



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> C12N 15/52, 15/63, 9/00, C07K 19/00, C12Q 1/68, A61K 38/43	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/07870</b> <b>(43) International Publication Date:</b> 26 February 1998 (26.02.98)
<b>(21) International Application Number:</b> PCT/US97/14093 <b>(22) International Filing Date:</b> 12 August 1997 (12.08.97)  <b>(30) Priority Data:</b> 08/703,623 23 August 1996 (23.08.96) US 08/816,462 13 March 1997 (13.03.97) US  <b>(71) Applicant:</b> BETH ISRAEL DEACONESS MEDICAL CENTER, INC. [US/US]; 330 Brookline Avenue, Boston, MA 02215 (US).  <b>(72) Inventors:</b> AVRAHAM, Shalom; 22 Heath Hill, Brookline, MA 02146 (US). AVRAHAM, Hava; 22 Heath Hill, Brookline, MA 02146 (US). GROOPMAN, Jerome, E.; 79 Druce Street, Brookline, MA 02146 (US).  <b>(74) Agents:</b> LOREN, Ralph, A. et al.; Lahive & Cockfield, L.L.P., 28 State Street, Boston, MA 02109 (US).		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> NOVEL <i>RAFTK</i> SIGNALING MOLECULES AND USES THEREFOR  <b>(57) Abstract</b>  The present invention relates to the discovery of novel " <i>RAFTK</i> " genes and polypeptides. Therapeutics, diagnostics and screening assays based on these molecules are also disclosed.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

## NOVEL *RAFTK* SIGNALING MOLECULES AND USES THEREFOR

### Government Funding

This work was supported, in part, by National Institutes of Health grants: HL 46668, HL 53745-02, HL 55187-01, HL 43510-06A, HL 43510-07, HL 51456-02 and HL 55445-01. The U.S. Government may, therefore, have certain rights to this invention.

### Background of the Invention

Signal transduction is triggered by stimulation of a cell surface receptor which either has kinase activity itself or is physically and/or functionally linked to an intracellular protein tyrosine kinase (PTK) (Cantley, L.C. et al. (1991) *Cell* 64, 281-302; Shattil, S.J., and Brugge, J.S. (1991) *Curr. Opin. Cell Biol.* 3, 869-879; Weiss, A. (1993) *Cell* 73, 209-212). PTKs participate in a variety of signal transduction pathways which modulate cell growth and differentiation (Ullrich, A., and Schlessinger, J. (1990) *Cell* 61, 203-212; Pawson, T., and Gish, G.D. (1992) *Cell* 71, 359-362; Fantl, W.J., et al. (1993) *Ann. Rev. Biochem.* 62, 453-481). Through a series of inducible and reversible protein-protein interactions and phosphorylation-mediated enzymatic activities, protein-tyrosine kinases are recruited to relay signals throughout the cell. Such interactions are involved in all stages of the intracellular signal transduction process - at the plasma membrane, where the signal is initiated; in the cytoplasm, where the signals are disseminated to different cellular locations; and in the nucleus, where other proteins involved in transcriptional control form complexes to regulate transcription of particular genes. Protein kinase cascades allow for amplification, feedback, cross-talk, and branching in signal transduction pathways.

The integrin cell surface receptors are also capable of transducing cytoplasmic signals (Hynes, R.O. (1992) *Cell* 69, 11-25; Juliano, R.L., and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585; Schwartz, M.A. (1992) *Trends Cell Biol.* 2, 304-308) and activation of this pathway is linked to one or more PTKs (Guan, J.-L., et al. (1991) *Cell Regul.* 2, 951-964; Kornberg, L.J. et al. (1991) *Proc. Natl. Acad. Sci. USA* 88, 8392-8396). Integrins are the major family of cell surface receptors that mediate adhesive interactions (Albelda, S. M. and Buck, C. A. (1990) *FASEB J* 4, 2868). Integrin receptor engagement and subsequent clustering of integrins lead to the formation of focal adhesion sites. Protein assemblies of integrins, linked to intracellular cytoskeletal complexes and to bundles of actin filaments in focal adhesions, play critical roles in modulating adhesion and inducing shape changes involved in cell spreading and locomotion (Hynes R O. (1992) *Cell* 69, 11). Such

cellular adhesive interactions, mediated by cell surface receptors that bind to ligands on adjacent cells or in the extracellular matrix, participate in the processes of cell migration, proliferation and differentiation (Gumbiner, B. M. (1993) *Neuron* 11, 551).

- 5           A large number of cytoplasmic proteins have been identified as components of focal adhesion structures (Burridge K, et al. (1998) *Ann Rev Cell Biol* 4, 487; Turner, C. E. and Burridge, K. (1991) *Curr Opin Cell Biol* 3, 849). These are classified as either structural proteins or signaling molecules. Vinculin,  $\alpha$ -actinin and talin are well known as the major structural proteins at focal adhesion sites (Turner, C. E. and
- 10   Burridge, K. (1991) *Curr Opin Cell Biol* 3, 849). *In vitro* and *in vivo* studies have shown that these proteins serve as bridge-like linkages between the integrins and actin filaments, and as a dock for the association of signaling proteins that leads to integrin-induced changes in cell function (Clark, E. A. and Brugge, J.S. (1995) *Science* 268, 233).
- 15           Several PTKs have been implicated in integrin signaling events by virtue of either their integrin-dependent activation or their localization to these focal contacts (Clark, E. A. and Brugge, J. S. (1995) *Science* 268, 233; Richardson, A. and Parsons, J. T. (1995) *Bioessays* 17, 229). Two focal adhesion proteins that demonstrate a high stoichiometry of tyrosine phosphorylation upon integrin activation are the focal
- 20   adhesion kinase (FAK) and paxillin (Schaller, M. D. et al. (1992) *Proc Natl Acad Sci USA* 89, 5192; Hanks, S. K. et al. (1992) *Proc Natl Acad Sci USA* 89, 8487; Burridge, K. et al. (1992) *J Cell Biol* 119, 893). The tyrosine phosphorylation of these two proteins has been suggested as being involved in both the formation of focal adhesions and the assembly of actin stress fibers (Burridge, K. et al. (1992) *J Cell Biol*
- 25   119, 893). In addition, the association of FAK with the cytoskeletal protein talin in NIH3T3 cells was observed (Chen, H-C et al. (1995) *J Biol Chem* 270, 16995).

pp125<sup>FAK</sup> is phosphorylated in response to  $\alpha$ IIb $\beta$ <sub>3</sub>- integrin-mediated cell adhesion (See, e.g., Juliano, R.L., and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585; Hanks, S.K. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 8487-8489; Schaller, M.D., et

30   al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196). Induction of the kinase activity and the tyrosine phosphorylation of pp125<sup>FAK</sup> were observed following the adherence of fibroblasts to fibronectin (See, e.g., Guan, J.-L. et al. (1991) *Cell Regul.* 2, 951-964; Kornberg, L.J. et al. (1991) *Proc. Natl. Acad. Sci. USA* 88, 8392-8396; Hanks, S.K. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 8487-8489; Schaller, M.D., et

35   al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196), the adherence of epidermal carcinoma cells to fibronectin, laminin, or collagen type IV (Kornberg, L., et al. (1992) *J. Biol. Chem.* 267, 23439-23442), and the aggregation of platelets in the



presence of fibrinogen, a ligand for  $\alpha 11b\beta_3$ - integrin (glycoprotein IIb/IIIa) (Lipfert, L., et al. (1992) *J. Cell Biol.* 119, 905-912). Phosphorylated pp125<sup>FAK</sup> is localized in focal adhesion contacts.

Src tyrosine kinases are capable of interacting with components of focal adhesions upon kinase activation (Weng, Z. et al. (1993) *J Biol Chem* 268, 14956; Schaller, M. D. et al. (1994) *Mol Cell Biol* 14, 1680). The Tyr<sup>397</sup> of FAK is autophosphorylated upon integrin-mediated stimulation. This phosphotyrosine then provides a binding site for the SH2 domain of p60Src and its family members, e.g. p59Fyn (Cobb, B. S. et al. (1994) *Mol Cell Biol* 14, 147 Eide, B. L. et al. (1995) *Mol Cell Biol* 15, 2819). Integrin-mediated signal transduction was found to be linked to the Ras pathway by Grb2 binding to FAK (Schlaepfer, D. D. et al. (1994) *Nature* 372, 786; Kharbanda, S. et al. (1995) *Proc Natl Acad Sci USA* 92, 6132). Phosphatidylinositol 3-kinase (PI-3 kinase) was also associated with the activated FAK (Chen, H-C et al. (1994) *Proc Natl Acad Sci USA* 91, 10148; Guinebault, C. et al. (1995) *J Cell Biol* 129, 831). In addition, FAK phosphorylation is stimulated by a number of other substances, including small peptide mitogens such as vasopressin, bombesin, endothelin (Zachary, I. et al. (1992) *J Biol Chem* 267, 19031; Sinnott-Smith, J. et al. (1993) *J Biol Chem* 268, 14261), and bradykinin (Leeb-Lundberg, L. M. F et al. (1994) *J Biol Chem* 269, 24328); bioactive lipids such as Alzheimer's Ab peptide (Zhang C et al. *J Biol Chem* 269, 25247, 1994); antigens for immunoglobulin E receptors (Hamawy, M. M. et al. (1993) *J Biol Chem* 268, 6851); neuropeptide receptors (Zhang, C. et al. (1994) *J. Biol. Chem.* 269, 25247-25250), growth factors such as hepatocyte growth factor, platelet-derived growth factor and M-CSF-1; (Kharbanda, S. et al. (1995) *Proc Natl Acad Sci USA* 92, 6132; Matsumoto, K. et al. (1994) *J Biol Chem* 269, 31807; Rankin, S. and Rozengurt, E. (1994) *J Biol Chem* 269, 704) and upon oncogenic transformation (Guan, J.-L., and Shalloway, D. (1992) *Nature* 358, 690-692) in adherent cells.

pp125<sup>FAK</sup> has been cloned from *Xenopus* (*X. Laevis*), avian, rodent, and human species and is expressed in a wide range of cell types (See, e.g., Schaller, M.D. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196; Schaller, M.D., and Parsons, J.T. (1993) *Trends Cell Biol.* 3, 258-262; Clark, E.A., and Brugge, J.S. (1995) *Science* 268, 233-239).

### Summary of the Invention

The present invention is based on the discovery of novel molecules, referred to herein as "related adhesion focal tyrosine kinase" or "RAFTK" polypeptide molecules and the nucleic acid molecules coding therefore. The RAFTK molecules of the

present invention are useful in regulating a variety of cellular processes. The *RAFTK* polypeptide is an intracytoplasmic protein tyrosine kinase.

In one aspect, the invention features isolated vertebrate *RAFTK* nucleic acid molecules. In a preferred embodiment a *RAFTK* nucleic acid has a nucleic acid  
5 sequence shown in one of SEQ ID NOs:1 or 3, or a complement or fragment thereof. The disclosed molecules can be non-coding, (e.g. probe, antisense or ribozyme molecules) or can encode a polypeptide with *RAFTK* bioactivity. In a preferred embodiment a *RAFTK* nucleic acid of the present invention comprises the coding region of one of SEQ ID NOs: 1 or 3. In another preferred embodiment the subject  
10 *RAFTK* nucleic acids encode a polypeptide with a *RAFTK* bioactivity. In a particularly preferred embodiment the nucleic acid of the present invention encodes a polypeptide shown in one of SEQ ID NOs: 2 or 4.

In one embodiment, the nucleic acids of the present invention can hybridize to a vertebrate *RAFTK* gene or to the complement of a vertebrate *RAFTK* gene. In a  
15 further embodiment, a *RAFTK* nucleic acid hybridizes with the coding sequence designated in one of SEQ ID NOs:1 or 3 or to the complement to the coding sequence designated in one of SEQ ID NOs:1 or 3. In a preferred embodiment, the hybridization is conducted under stringent conditions.

In further embodiments, the nucleic acid molecule is a *RAFTK* nucleic acid  
20 molecule that is at least 60%, at least 70%, preferably 80%, more preferably 85%, and even more preferably at least 95% homologous in sequence to the nucleic acids shown in one of SEQ ID NOs:1 or 3 or to the complement of the nucleic acid shown in one of SEQ ID NOs:1 or 3. In another embodiment, the *RAFTK* nucleic acid molecule encodes a polypeptide that is at least 60%, preferably at least 70%, preferably 80%,  
25 and more preferably at least 85%, and even more preferably at least 95% homologous in sequence to the polypeptide shown in one of SEQ ID NOs: 2 or 4.

The invention also provides probes and primers comprising substantially purified oligonucleotides, which correspond to a region of nucleotide sequence which hybridizes to at least 6 consecutive nucleotides of the sequence set forth in one of  
30 SEQ ID NOs:1 or 3, the complement of one of SEQ ID NOs:1 or 3, or naturally occurring mutants thereof. In a preferred embodiment a probe or primer of the present invention hybridizes under stringent conditions to a nucleic acid corresponding to at least 12 consecutive nucleotides of either sense or antisense sequence of one or more of SEQ ID NOs:1 or 3; preferably to at least 25 consecutive nucleotides; and more  
35 preferably to at least 40, 50 or 75 consecutive nucleotides of either sense or antisense sequence of one or more of SEQ ID NOs:1 or 3.

In a preferred embodiment, a probe of the present invention comprises all or a portion of nucleotides 1595-2974 of one of SEQ ID NOs:1 or 3. In preferred embodiments, the probe/primer further includes a label group, which is capable of being detected.

5 For expression, the subject *RAFTK* nucleic acids can include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter (e.g., for constitutive expression or inducible expression) or transcriptional enhancer sequence, which regulatory sequence is operably linked to the *RAFTK* gene sequence. Such regulatory sequences in conjunction with a *RAFTK* nucleic acid molecules can be  
10 useful vectors for gene expression. This invention also features host cells transfected with such an expression vector whether prokaryotic or eukaryotic and *in vitro* (e.g. cell culture) and *in vivo* (e.g. transgenic) methods for producing *RAFTK* polypeptides by employing the expression vectors.

The invention also features transgenic non-human animals which include a  
15 heterologous form of a *RAFTK* gene described herein, or which misexpress an endogenous *RAFTK* gene (e.g., an animal in which expression of one or more of the subject *RAFTK* proteins is disrupted). Such a transgenic animal can serve as an animal model for studying cellular and tissue disorders comprising mutated or mis-expressed *RAFTK* alleles or can be used in drug screening. Alternatively, such a  
20 transgenic animal can be useful for expressing recombinant *RAFTK* polypeptides.

In another aspect, the invention features isolated *RAFTK* polypeptides, preferably substantially pure preparations e.g., of plasma purified or recombinantly produced *RAFTK* polypeptides. In preferred embodiments, the polypeptide has a *RAFTK* bioactivity. In addition, *RAFTK* polypeptides which specifically antagonize  
25 the activity of a native *RAFTK* polypeptide, such as can be provided by truncation mutants or other dominant negative mutants, are also specifically contemplated by the present invention.

In one embodiment, the polypeptide is identical to or homologous to a *RAFTK* protein represented in one of SEQ ID NOs: 2 or 4. Related members of the vertebrate  
30 and particularly the mammalian *RAFTK* family are also within the scope of the invention. Preferably, a *RAFTK* polypeptide has an amino acid sequence at least 60%, at least 70% homologous, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% homologous to the polypeptide represented by one of SEQ ID NOs: 2 or 4. In a preferred embodiment, the *RAFTK* polypeptide is  
35 encoded by a nucleic acid which hybridizes with a nucleic acid sequence represented in one of SEQ ID NOs: 1 or 3. The subject *RAFTK* polypeptides also include modified polypeptides, which are resistant to post-translation modification, as for

example, due to mutations which alter modification sites (such as tyrosine, threonine, serine or asparagine residues), or which prevent glycosylation of the protein, or which prevent interaction of the protein with intracellular proteins.

The *RAFTK* polypeptide can comprise a full length protein, such as  
5 represented in one of SEQ ID NOs: 2 or 4, or it can comprise a fragment  
corresponding to one or more particular motifs/domains, or to arbitrary sizes, e.g., at  
least 5, 10, 25, 50, 100, 150 or 200 amino acids in length. In preferred embodiments,  
the *RAFTK* polypeptide includes at least a portion of an *RAFTK* kinase domain and  
has a *RAFTK* bioactivity. In preferred embodiments the subject *RAFTK* polypeptides  
10 have a molecular weight of approximately 123kD.

Another aspect of the invention features chimeric molecules (e.g. fusion  
proteins) comprised of a *RAFTK* polypeptides. For instance, the *RAFTK* polypeptides  
can be provided as a recombinant fusion protein which includes a second polypeptide  
portion, e.g., a second polypeptide having an amino acid sequence unrelated  
15 (heterologous) to the *RAFTK* polypeptide, (e.g. the second polypeptide portion is  
glutathione-S-transferase, an enzymatic activity such as alkaline phosphatase or an  
epitope tag).

A further aspect of the invention features pharmaceutical preparations  
including *RAFTK* polypeptides or homologues, or the nucleic acids encoding *RAFTK*  
20 polypeptides and a pharmaceutically acceptable carrier.

Yet another aspect of the present invention pertains to an immunogen  
comprising a *RAFTK* polypeptide in an immunogenic preparation, the immunogen  
being capable of eliciting an immune response specific for a *RAFTK* polypeptide,  
e.g., a humoral response, an antibody response and/or cellular response. In preferred  
25 embodiments, the immunogen comprises an antigenic determinant, e.g. a unique  
determinant, from the protein represented by one of SEQ ID NOs: 2 or 4.

A still further aspect of the present invention features antibodies and antibody  
preparations specifically reactive with an epitope of the *RAFTK* protein. In preferred  
embodiments the antibody specifically binds to an epitope represented in one of SEQ  
30 ID NOs:2 or 4. In a particularly preferred embodiment, an antibody of the present  
specifically recognizes amino acids 68-1009 from the *RAFTK* c-terminus.

Yet another aspect of the present invention concerns a method for modulating  
the growth, migration, differentiation, and/or survival of a cell, e.g., a mast cell, a  
melanocyte, or a megakaryocyte, by modulating *RAFTK* bioactivity (e.g., by  
35 potentiating or disrupting certain protein-protein interactions in a *RAFTK* signaling  
pathway). In general, whether carried out *in vivo*, *in vitro*, or *in situ*, the method  
comprises treating the cell with an effective amount of a *RAFTK* therapeutic so as to

alter, relative to the cell in the absence of treatment, at least one of (i) rate of growth, (ii) differentiation, (iii) hematopoiesis or (iv) survival of the cell. In preferred embodiments the cells are selected from a group including mast cells, melanocytes, and megakaryocytic cells. In another embodiment a *RAFTK* therapeutic can be used in a method of modulating cell adhesion, migration, phagocytosis, or motility. In preferred embodiments, the method can be used to modulate focal adhesion formation or to treat metastasis by a tumor cell.

Accordingly, the method can be carried out with *RAFTK* therapeutics such as peptide and peptidomimetics or other molecules identified in the above-referenced drug screens which agonize or antagonize the effects of signaling from a *RAFTK* protein or ligand binding of a *RAFTK* protein. Other *RAFTK* therapeutics include antisense constructs for inhibiting expression of *RAFTK* proteins, and dominant negative mutants of *RAFTK* proteins which competitively inhibit ligand interactions upstream and signal transduction downstream of the wild-type *RAFTK* protein. In a preferred embodiment the subject *RAFTK* peptides are capable of modulating signal transduction in a pathway involving stem cell factor, thrombin, fibronectin, CSF-1/M-CSF, T cell receptor stimulation, bFGF, oncoprotein M, IL-6, or TNF $\alpha$ .

A further aspect of the present invention provides a method of determining if a subject is at risk for a disorder characterized by unwanted cell proliferation or aberrant control of differentiation. The method includes detecting, in a tissue of the subject, the presence or absence of a genetic lesion characterized by at least one of (i) a mutation of a gene encoding a *RAFTK* protein, e.g. represented in one of SEQ ID NOs: 1 or 3, or a homolog thereof; or (ii) the mis-expression of a *RAFTK* gene. In preferred embodiments, detecting the genetic lesion includes ascertaining the existence of at least one of: a deletion of one or more nucleotides from a *RAFTK* gene; an addition of one or more nucleotides to the gene, a substitution of one or more nucleotides of the gene, a gross chromosomal rearrangement of the gene; an alteration in the level of a messenger RNA transcript of the gene; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; a non-wild type level of the protein; and/or an aberrant level of *RAFTK* protein.

For example, detecting the genetic lesion can include (i) providing a probe/primer comprised of an oligonucleotide which hybridizes to a sense or antisense sequence of a *RAFTK* gene or naturally occurring mutants thereof, or 5' or 3' flanking sequences naturally associated with the *RAFTK* gene; (ii) contacting the probe/primer with an appropriate nucleic acid containing sample; and (iii) detecting, by hybridization of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion; e.g., wherein detecting the lesion comprises utilizing the probe/primer

to determine the nucleotide sequence of the *RAFTK* gene and, optionally, of the flanking nucleic acid sequences. For example, the primer can be employed in a polymerase chain reaction (PCR) or in a ligation chain reaction (LCR). In alternate embodiments, the level of a *RAFTK* protein is detected in an immunoassay using an antibody which is specifically immunoreactive with the *RAFTK* protein.

In a further aspect the invention provides for methods of preparing differentiated blood cells by modulating the activity of a *RAFTK* protein in a progenitor stem cell. In a preferred embodiment the subject method can be used to prepare megakaryocytes. Platelets can also be prepared using the subject method.

In yet another aspect, the invention provides assays, e.g., for screening test compounds to identify inhibitors, or alternatively, potentiators, of an interaction between a *RAFTK* protein and, for example, an intracellular protein which binds to the *RAFTK* protein. An exemplary method includes the steps of (i) combining a *RAFTK* polypeptide or bioactive fragments thereof, a *RAFTK* target molecule, and a test compound, e.g., under conditions wherein, but for the test compound, the *RAFTK* protein and target molecule are able to interact; and (ii) detecting the formation of a complex which includes the *RAFTK* protein and the target polypeptide either by directly quantitating the complex, or by measuring a bioactivity of the *RAFTK* protein. Several *RAFTK* binding-proteins have been identified and any of these novel interactions can be exploited in the subject drug screening assays. A statistically significant change, such as a decrease, in the interaction of the *RAFTK* and target molecule in the presence of a test compound (relative to what is detected in the absence of the test compound) is indicative of a modulation (e.g., inhibition or potentiation of the interaction between the *RAFTK* protein and the target molecule). In preferred embodiments the ability of a compound to affect the interaction between and one or more of the *RAFTK* binding-proteins selected from the group consisting of paxillin, protein kinase C (PKC)- $\alpha$ , PKC- $\delta$ , src, fyn, Grb2, PI3 kinase, the c-fms receptor, and calpain, is detected. In certain embodiments the phosphorylation state of *RAFTK* or a *RAFTK* binding protein is measured as a readout of protein-protein interaction. In certain embodiments the reaction mixture can be a reconstituted protein mixture or a cell lysate. In certain embodiments the *RAFTK* protein can be a recombinant protein. In certain embodiments either the *RAFTK* protein or the *RAFTK*-binding protein is a fusion protein and in preferred embodiments, at least one of the proteins includes a label group for detection. In yet another embodiment the reaction mixture is a whole cell and the interaction of *RAFTK* and a *RAFTK* binding protein is detected in a two hybrid assay.

In yet another aspect the invention provides for compounds identified using the subject assay, whether agonists or antagonists (inhibitors) of *RAFTK* activity. In one embodiment the compounds identified in the subject screening assays are included in a pharmaceutical preparation. In yet another embodiment the invention provides for a method of modulating cell growth, differentiation or survival by contacting a cell with a pharmaceutical preparation including a compound identified in one of the subject drug screening assays.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

### Brief Description of the Drawings

*Figure 1* is a schematic representation and restriction enzyme map of the *RAFTK* cDNA. The various cDNA clones, obtained from the Human Hippocampus cDNA Library (in Zap II vector) and the CMK/PMA cDNA library (in  $\lambda$ -gt10 vector) are shown as indicated. Restriction enzyme sites are indicated along the length of the cDNA.

*Figure 2* shows a comparison of the deduced amino acid sequence of *RAFTK* with those of m-pp125<sup>FAK</sup>, src, c-fyn, htk and fgfr. Gaps (indicated by dashes) are introduced to optimize the alignment. Amino acid residues found to be conserved are boxed.

*Figure 3* shows an alignment of the predicted amino acid sequences (single-letter code) of the mouse *RAFTK*, human *RAFTK* and the mouse pp125<sup>FAK</sup> gene translated product. Amino acid residues found to be conserved are boxed.

*Figure 4* shows haplotype analysis of Chromosome 14 genetic markers in (C57BL/6J) x M. spretus)F<sub>1</sub> x M. spretus (BSS type) backcross mice showing linkage and relative position of *RAFTK*. Closed boxes indicate the inheritance of the C57BL/6J (B) allele and open boxes indicate the inheritance of the M. spretus (S) allele from the (C57BL/6J) x M. spretus)F<sub>1</sub> parent. Gene names and references to these loci can be found in GBASE. The first two columns indicate the number of backcross progeny with no recombinations. The following columns indicate recombinational events between adjacent loci (signified by a change from open box to closed box). The number of recombinants are listed below each column and the recombination frequency (REC %) between adjacent loci is indicated.

*Figure 5* shows co-segregation of *RAFTK* and Gnrh in BXD RI lines and localization to Chromosome 14. Strain distribution patterns are depicted for *RAFTK* in the BXD RI lines. The RI line distribution pattern is compared with that of the Gnrh

locus. Map units are indicated between *RAFTK* and *Gnrh*, as are 95% confidence limits.

### Detailed Description of the Invention

5 Protein tyrosine kinases (PTKs) play salient roles in a variety of signal transduction pathways which modulate cell growth and differentiation (Ullrich, A., and Schlessinger, J. (1990) *Cell* 61, 203-212; Pawson, T., and Gish, G.D. (1992) *Cell* 71, 359-362; Fantl, W.J., et al. (1993) *Ann. Rev. Biochem.* 62, 453-481). The novel *RAFTK* proteins of the present invention were identified using PCR primers based on  
10 conserved sequences of protein-tyrosine kinases.

The human *RAFTK* gene was cloned from the CMK cell line, which has properties of cells of the megakaryocytic lineage. The 3.6 kb *RAFTK* human cDNA is shown in SEQ ID NO:1. The full length cDNA contains an open reading frame with the first in-frame ATG codon located at nucleotides 294-296, followed by a stop  
15 codon at positions 3260-3262. *RAFTK* is 51% homologous to focal adhesion kinase, FAK at the nucleic acid level.

The murine homolog was subsequently cloned, based on the ability of a probe derived from the human sequence to hybridize to the mouse gene under high stringency conditions. The murine *RAFTK* homolog encodes a protein of 1009 amino  
20 acids. The amino acid sequences of the human and murine *RAFTK* proteins are 95% homologous and the nucleic acid sequences are 90% homologous. The *RAFTK* gene was mapped to human chromosome 8 and to chromosome 14 in the mouse.

The open reading frame of the *RAFTK* nucleic acid encodes a predicted protein of 1009 amino acid residues. The *RAFTK* protein migrates with a molecular  
25 weight of approximately 115-125 kD. In preferred embodiments, the *RAFTK* protein of the present invention is approximately about 123 kD. The *RAFTK* proteins of the present invention can be activated by phosphorylation, and it will be understood that other post-translational modifications can alter the apparent molecular weight of the protein.

30 The *RAFTK* polypeptide contains several structural motifs common to all protein kinases, including the putative ATP-binding site (G<sup>432</sup>-X-G<sup>434</sup>-X-X-G<sup>437</sup>, where X is any amino acid) and three residues that are predicted to interact with the  $\gamma$  phosphate group of the bound ATP molecule (at amino acids 402, 529, and 655). The amino acid sequences at positions 549-554 and 588-592 are also conserved among  
35 protein-tyrosine kinases. The kinase domain consists primarily of the catalytic domain including the putative ATP-binding site (amino acids 432-437). Amino acid residues 880-887 are important in mediating association of *RAFTK* with src and/or



fyn, particularly when the tyrosine residue at 881 is phosphorylated. The kinase domain of *RAFTK* is flanked by N-terminal and C-terminal regions; the N-terminal region of the protein (amino acids 1-39) is unique and the C-terminal region contains a proline-rich stretch (residues 690-767) in which the proline content exceeds 20%.

5 *RAFTK* lacks myristilation sites and SH2 and SH3 domains.

*RAFTK* was found to be expressed in fetal brain, lung, and liver, and to have a less restricted pattern of expression in adults. Expression was detected in a variety of adult tissues, including: CD34+ bone marrow cells, megakaryocytes, platelets, brain (particularly in the amygdala and hippocampus), macrophages, peripheral blood  
10 lymphocytes, spleen, thymus, B lymphocytes, T lymphocytes, and certain cancer cells.

Accordingly, certain aspects of the present invention relate to nucleic acid molecules encoding a vertebrate, e.g., mammalian *RAFTK* polypeptides, the *RAFTK* polypeptides, antibodies immunoreactive with *RAFTK* polypeptides, and preparations  
15 of such compositions. In addition, drug discovery assays are provided for identifying agents which can modulate the biological function of *RAFTK* proteins. Such agents can be useful therapeutically, therefore, to alter the growth and/or differentiation of a cell. Moreover, the present invention provides diagnostic and therapeutic assays and reagents for detecting and treating disorders involving, for example, aberrant  
20 expression (or loss thereof) of mammalian *RAFTK* genes. Other aspects of the invention are described below or will be apparent to those skilled in the art in light of the present disclosure.

### Definitions

25 For convenience, the meaning of certain terms and phrases employed in the specification, examples, and appended claims are provided below.

The term "binding" as used herein is meant to include detectable interactions between molecules, such as can be detected using, for example, a yeast two hybrid  
30 assay. Also encompassed by this term are enzyme/substrate interactions (e.g. phosphorylation). Interactions between a *RAFTK* protein and a *RAFTK*-binding protein can be constitutive, or induced upon stimulation of a cell.

The term "bioactivity" of a *RAFTK* protein is intended to include effects on growth, differentiation, survival, and motility, e.g., migration or adhesion of cells.  
35 *RAFTK* has been shown to have broad involvement in numerous signaling pathways, and to be activated by: stem cell factor, thrombin stimulation, fibronectin, CSF-1/M-CSF, T cell receptor stimulation, bFGF, oncoprotein M, IL-6, and TNF $\alpha$ . *RAFTK* is

also activated by changes intracellular calcium levels, and by activation of protein kinases  $\alpha$  and  $\delta$ . Thus, *RAFTK* is capable of modulating the growth, differentiation, survival, and motility of numerous cell types, including megakaryocytes, T cells, B cells, monocytes, hematopoietic stem cells (e.g., CD34<sup>+</sup> bone marrow cells),  
5 melanocytes, neural cells (particularly in the amygdala and hippocampus), macrophages, peripheral blood lymphocytes, spleen, thymus, B lymphocytes, T lymphocytes, and certain cancer cells (e.g., Kaposi's sarcoma cells). The subject *RAFTK* polypeptides are also capable of modulating platelet function.

*RAFTK* also modulates the formation of focal adhesions and actin stress fibers,  
10 and is thus important in the control of metastatic growth and in the normal cell growth and integrity, and in processes which involve cell motility, such as, for example, phagocytosis. Other bioactivities of the subject *RAFTK* polypeptides are described in more detail herein.

"Cells," "host cells" or "recombinant host cells" are terms used  
15 interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications can occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

20 A "chimeric protein" or "fusion protein" is a fusion of a first amino acid sequence encoding one of the subject mammalian *RAFTK* polypeptides with a second amino acid sequence defining a domain (e.g. polypeptide portion) foreign to and not substantially homologous with any domain of one of the mammalian *RAFTK* polypeptides. A chimeric protein may present a foreign domain which is found (albeit  
25 in a different protein) in an organism which also expresses the first protein, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of organisms. In general, a fusion protein can be represented by the general formula X-*RAFTK*-Y, wherein *RAFTK* represents a portion of the protein which is derived from one of the mammalian *RAFTK* proteins, and X and Y are  
30 independently absent or represent amino acid sequences which are not related to one of the mammalian *RAFTK* sequences in an organism, including naturally occurring mutants.

"Complementary" sequences as used herein refer to sequences which have  
sufficient complementarity to be able to hybridize, forming a stable duplex.

35 A "delivery complex" as used herein refers to a targeting means (e.g. a molecule that results in higher affinity binding of a gene, protein, polypeptide or peptide to a target cell surface and/or increased cellular uptake by a target cell).

Examples of targeting means include: sterols (e.g. cholesterol), lipids (e.g. a cationic lipid, virosome or liposome), viruses (e.g. adenovirus, adeno-associated virus, and retrovirus) or target cell specific binding agents (e.g. ligands recognized by target cell specific receptors).

5           As is well known, genes for a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity. The term "DNA sequence encoding a mammalian *RAFTK*

10 polypeptide" can refer to one or more genes within a particular individual. Moreover, certain differences in nucleotide sequences can exist between individual organisms, which are called alleles. Such allelic differences can result in differences in amino acid sequence of the encoded polypeptide yet still encode a protein with the same biological activity.

15           As used herein, the term "gene" or "recombinant gene" refers to a nucleic acid molecule comprising an open reading frame encoding one of the mammalian *RAFTK* polypeptides of the present invention, including both exon and (optionally) intron sequences. A "recombinant gene" refers to nucleic acid encoding a mammalian *RAFTK* polypeptide and comprising mammalian *RAFTK*-encoding exon sequences,

20 though it may optionally include intron sequences which are either derived from a chromosomal mammalian *RAFTK* gene or from an unrelated chromosomal gene. Exemplary recombinant genes encoding the subject mammalian *RAFTK* polypeptides are represented in the appended Sequence Listing. The term "intron" refers to a DNA sequence present in a given mammalian *RAFTK* gene which is not translated into

25 protein and is generally found between exons.

"Homology" or "identity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then

30 the molecules are homologous or identical at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. An "unrelated" or "non-homologous" sequence shares less than 40% identity, though preferably less than 25% identity, with one of the mammalian *RAFTK* sequences of the present invention.

35           The term "isolated" as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs, or RNAs, respectively, that are present in the natural source of the nucleic acids. For example, an isolated nucleic

acid encoding one of the subject mammalian *RAFTK* polypeptides preferably includes no more than 10 kilobases (kb) of nucleic acid sequence which naturally immediately flanks the mammalian *RAFTK* gene in genomic DNA, more preferably no more than 5kb of such naturally occurring flanking sequences, and most preferably less than 1.5kb of such naturally occurring flanking sequence. Moreover, an "isolated" nucleic acid" includes nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state. The term "isolated" as used herein also refers to a polypeptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state. The term "isolated" is also used herein to refer to polypeptides which are isolated from other cellular proteins and is meant to encompass both purified and recombinant polypeptides.

The term "modulation" as used herein refers to both upregulation, i.e., stimulation or potentiation, and downregulation, i.e. suppression, of a response.

The "non-human animals" of the invention include mammals such as rodents, non-human primates, sheep, dog, cow, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the rodent family including rat and mouse, most preferably mouse, though transgenic amphibians, such as members of the *Xenopus* genus, and transgenic chickens are also contemplated herein. The term "chimeric animal" is used herein to refer to animals in which the recombinant gene is found, or in which the recombinant is expressed in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that one of the recombinant mammalian *RAFTK* genes is present and/or expressed or disrupted in some tissues but not others.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides.

The terms "protein", "polypeptide" and "peptide" are used interchangeably herein.

The term "recombinant protein" refers to a polypeptide of the present invention which is produced by recombinant DNA techniques, wherein generally, DNA encoding a mammalian *RAFTK* polypeptide is inserted into a suitable

expression vector which is in turn used to transform a host cell to produce the heterologous protein. Moreover, the phrase "derived from", with respect to a recombinant *RAFTK* gene, is meant to include within the meaning of "recombinant protein" those proteins having an amino acid sequence of a native *RAFTK* protein, or  
5 an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions (including truncation) of a naturally occurring form of the protein.

As used herein, the term "specifically hybridizes" or "specifically detects" refers to the ability of a nucleic acid molecule of the invention to hybridize to at least  
10 approximately 12, 20, 30, 50, 100, 150, 200, or 300 consecutive nucleotides of a vertebrate, preferably mammalian, *RAFTK* gene, such as a *RAFTK* sequence designated in one of SEQ ID NOs:1 or 3, or a sequence complementary thereto, or naturally occurring mutants thereof, such that it shows more than 10 times more hybridization, preferably more than 100 times more hybridization, and even more  
15 preferably more than 100 times more hybridization than it does to to a cellular nucleic acid (e.g., mRNA or genomic DNA) encoding a protein other than a vertebrate, preferably mammalian, *RAFTK* protein as defined herein.

As used herein, the term "tissue-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence  
20 operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue, such as cells of hepatic or pancreatic origin, neuronal cells, or immune cells. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

"Transcriptional regulatory sequence" is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of one of the recombinant mammalian *RAFTK* genes is under the control of a promoter sequence  
30 (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring forms of *RAFTK* proteins.

35 As used herein, the term "transfection" means the introduction of a nucleic acid, e.g., an expression vector, into a recipient cell by nucleic acid-mediated gene transfer. "Transformation", as used herein, refers to a process in which a cell's

genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of a mammalian *RAFTK* polypeptide or, in the case of anti-sense expression from the transferred gene, the expression of a naturally-occurring form of the *RAFTK* protein is disrupted.

5 As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., one of the mammalian *RAFTK* polypeptides, or an transcript which is antisense to a *RAFTK* nucleic acid sequence), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but  
10 which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal  
15 expression of a selected nucleic acid.

A "transgenic animal" refers to any animal, preferably a non-human mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the  
20 cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule can be integrated within a chromosome, or it can be extrachromosomally replicating DNA. In the typical transgenic animals  
25 described herein, the transgene causes cells to express a recombinant form of one of the mammalian *RAFTK* proteins, e.g. either agonistic or antagonistic forms. However, transgenic animals in which the recombinant *RAFTK* gene is silent are also contemplated, as for example, the FLP or CRE recombinase dependent constructs  
30 described below. Moreover, "transgenic animal" also includes those recombinant animals in which gene disruption of one or more *RAFTK* genes is caused by human intervention, including both recombination and antisense techniques.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. A preferred vector is an  
35 episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which

they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specification, 5 "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

#### 10 *Nucleic Acids of the Present Invention*

As described below, one aspect of the invention pertains to isolated nucleic acids comprising nucleotide sequences encoding *RAFTK* polypeptides, and/or equivalents of such nucleic acids. The term "equivalent" includes nucleotide sequences encoding functionally equivalent *RAFTK* polypeptides or functionally 15 equivalent peptides having a bioactivity of a vertebrate *RAFTK* protein such as described herein. Equivalent nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and, therefore, include sequences that differ from the nucleotide sequence of the *RAFTK* gene shown in one of SEQ ID NOs:1 or 3 due to the degeneracy of the 20 genetic code.

Preferred nucleic acids are vertebrate *RAFTK* nucleic acids. Particularly preferred vertebrate *RAFTK* nucleic acids are mammalian. Regardless of species, *RAFTK* nucleic acids encode polypeptides that are at least 60% similar to an amino acid sequence of a vertebrate *RAFTK*. Preferred nucleic acids encode a *RAFTK* 25 polypeptide comprising an amino acid sequence at least 60%, at least 70% homologous, preferably at least 80% homologous, more preferably at least 90% homologous with an amino acid sequence of a vertebrate *RAFTK*, e.g., such as a sequence shown in one of SEQ ID NOs:2 or 4. Nucleic acids which encode polypeptides at least about 95%, and even more preferably at least about 98-99% 30 similarity with an amino acid sequence represented in one of SEQ ID NOs:2 or 4 are most preferred. Still other preferred nucleic acids of the present invention encode a *RAFTK* polypeptide which includes a polypeptide sequence corresponding to all or a portion of amino acid residues of one of SEQ ID NOs:2 or 4, e.g., at least 5, 10, 25, 50, 100, 150 or 200 amino acid residues of that region.

35 Another aspect of the invention provides a nucleic acid which hybridizes under stringent conditions to a nucleic acid represented by one of SEQ ID NOs:1 or 3. Appropriate stringency conditions which promote DNA hybridization, for example,

6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C, are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50°C to a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or temperature of salt concentration may be held constant while the other variable is changed. In a particularly preferred embodiment, a *RAFTK* nucleic acid of the present invention binds to one of SEQ ID NOs: 1 or 3 under stringent conditions.

Preferred nucleic acids have a sequence at least 60%, at least 70% homologous and more preferably 80% and even more preferably at least 85% homologous with an amino acid sequence of a mammalian *RAFTK*, e.g., such as a sequence shown in SEQ ID NOs: 1. Nucleic acids at least 90%, more preferably 95%, and most preferably at least about 98-99% homologous with a nucleic sequence represented in SEQ ID NOs: 1 are of course also within the scope of the invention. In preferred embodiments, the nucleic acid is a mammalian *RAFTK* gene and in particularly preferred embodiments, includes all or a portion of the nucleotide sequence corresponding to the coding region of SEQ ID NOs: 1 or 3.

In preferred embodiments, the nucleic acid is a cDNA encoding a polypeptide having at least one bioactivity of the subject *RAFTK* polypeptide.

Nucleic acids having a sequence that differs from the nucleotide sequences shown in one of SEQ ID NOs: 1 or 3 due to degeneracy in the genetic code are also within the scope of the invention. Such nucleic acids encode functionally equivalent peptides (i.e., a peptide having a biological activity of a mammalian *RAFTK* polypeptide) but differ in sequence from the sequence shown in the sequence listing due to degeneracy in the genetic code. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC each encode histidine) may result in "silent" mutations which do not affect the amino acid sequence of a mammalian *RAFTK* polypeptide. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject *RAFTK* polypeptides will exist among mammals. One skilled in the art will appreciate that these variations in one or more nucleotides (e.g., up to about 3-5% of the nucleotides) of the nucleic acids encoding polypeptides having an activity of a mammalian *RAFTK* polypeptide can exist among individuals of a given species due to natural allelic variation.



As indicated by the examples set out below, *RAFTK* protein-encoding nucleic acids can be obtained from mRNA present in any of a number of eukaryotic cells. Nucleic acids encoding mammalian *RAFTK* polypeptides of the present invention can also be obtained from genomic DNA from both adults and embryos. For example, a

5 gene encoding a *RAFTK* protein can be cloned from either a cDNA or a genomic library in accordance with protocols described herein, as well as those generally known to persons skilled in the art. Examples of tissues and/or libraries suitable for isolation of the subject nucleic acids include brain, thymus, spleen, among others. A cDNA encoding a *RAFTK* protein can be obtained by isolating total mRNA from a

10 cell, e.g. a vertebrate cell, a mammalian cell, or a human cell, including embryonic cells. Double stranded cDNAs can then be prepared from the total mRNA, and subsequently inserted into a suitable plasmid or bacteriophage vector using any one of a number of known techniques. The gene encoding a mammalian *RAFTK* protein can also be cloned using established polymerase chain reaction techniques in accordance

15 with the nucleotide sequence information provided by the invention. The nucleic acid of the invention can be DNA or RNA. A preferred nucleic acid is a cDNA represented by a sequence shown in one of SEQ ID NOs:1 or 3.

#### *Vectors.*

20 This invention also provides expression vectors containing a nucleic acid encoding a *RAFTK* polypeptide, operably linked to at least one transcriptional regulatory sequence. "Operably linked" is intended to mean that the nucleotide sequence is linked to a regulatory sequence in a manner which allows expression of the nucleotide sequence. Regulatory sequences are art-recognized and are selected to

25 direct expression of the subject mammalian *RAFTK* proteins. Accordingly, the term "transcriptional regulatory sequence" includes promoters, enhancers and other expression control elements. Such regulatory sequences are described in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). In one embodiment, the expression vector includes a recombinant

30 gene encoding a peptide having an agonistic activity of a subject *RAFTK* polypeptide, or alternatively, encoding a peptide which is an antagonistic form of the *RAFTK* protein. Such expression vectors can be used to transfect cells and thereby produce polypeptides, including fusion proteins, encoded by nucleic acids as described herein. Moreover, the gene constructs of the present invention can also be used as a part of a

35 gene therapy protocol to deliver nucleic acids encoding either an agonistic or antagonistic form of one of the subject mammalian *RAFTK* proteins. Thus, another aspect of the invention features expression vectors for *in vivo* or *in vitro* transfection

and expression of a mammalian *RAFTK* polypeptide in particular cell types so as to reconstitute the function of, or alternatively, abrogate the function of *RAFTK* in a tissue. This is desirable, for example, when the naturally-occurring form of the protein is misexpressed; or to deliver a form of the protein which alters survival of tissue. Expression vectors can also be employed to inhibit neoplastic transformation.

In addition to viral transfer methods, such as those illustrated above, non-viral methods can also be employed to cause expression of a subject *RAFTK* polypeptide in the tissue of an animal. Most nonviral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral targeting means of the present invention rely on endocytic pathways for the uptake of the subject *RAFTK* polypeptide gene by the targeted cell. Exemplary targeting means of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes.

#### *Probes and Primers*

Moreover, the nucleotide sequences determined from the cloning of *RAFTK* genes from mammalian organisms allow for the generation of probes and primers designed for use in identifying and/or cloning *RAFTK* homologs in other cell types, e.g. from other tissues, as well as *RAFTK* homologs from other mammalian organisms. For instance, the present invention also provides a probe/primer comprising a substantially purified oligonucleotide, which oligonucleotide comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least approximately 12, preferably 25, more preferably 40, 50 or 75 consecutive nucleotides of sense or anti-sense sequence of one of SEQ ID NOs:1 or 3, or naturally occurring mutants thereof. For instance, primers based on the nucleic acid represented in one of SEQ ID NOs:1 or 3 can be used in PCR reactions to clone *RAFTK* homologs.

Likewise, probes based on the subject *RAFTK* sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto and able to be detected, e.g. the label group is selected from amongst radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors.

As discussed in more detail below, such probes can also be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a *RAFTK* protein, such as by measuring a level of a *RAFTK*-encoding nucleic acid in a sample of cells from a patient; e.g. detecting *RAFTK* mRNA levels or determining whether a genomic *RAFTK* gene has been mutated or deleted. Briefly, nucleotide probes can be generated from the subject *RAFTK* genes which facilitate histological screening of

intact tissue and tissue samples for the presence (or absence) of *RAFTK*-encoding transcripts. Similar to the diagnostic uses of anti-*RAFTK* antibodies, the use of probes directed to *RAFTK* messages, or to genomic *RAFTK* sequences, can be used for both predictive and therapeutic evaluation of allelic mutations which might be manifest in, for example, degenerative, neoplastic or hyperplastic disorders (e.g. unwanted cell growth) or abnormal differentiation of tissue. Used in conjunction with immunoassays as described herein, the oligonucleotide probes can help facilitate the determination of the molecular basis for a disorder which may involve some abnormality associated with expression (or lack thereof) of a *RAFTK* protein. For instance, variation in polypeptide synthesis can be differentiated from a mutation in a coding sequence.

#### *Antisense, Ribozyme and Triplex techniques*

One aspect of the invention relates to the use of the isolated nucleic acid in "antisense" therapy. As used herein, "antisense" therapy refers to administration or *in situ* generation of oligonucleotide molecules or their derivatives which specifically hybridize (e.g. bind) under cellular conditions, with the cellular mRNA and/or genomic DNA encoding one or more of the subject *RAFTK* proteins so as to inhibit expression of that protein, e.g. by inhibiting transcription and/or translation. The binding can be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, "antisense" therapy refers to the range of techniques generally employed in the art, and includes any therapy which relies on specific binding to oligonucleotide sequences.

An antisense construct of the present invention can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes a mammalian *RAFTK* protein. Alternatively, the antisense construct is an oligonucleotide probe which is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences of a mammalian *RAFTK* gene. Such oligonucleotide probes are preferably modified oligonucleotides which are resistant to endogenous nucleases, e.g. exonucleases and/or endonucleases, and are therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example,

by Van der Krol et al. (1988) *Biotechniques* 6, 958-976; and Stein et al. (1988) *Cancer Res* 48, 2659-2668.

Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to *RAFTK* mRNA. The antisense oligonucleotides  
5 bind to the *RAFTK* mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double-stranded antisense nucleic acids, a single strand of the duplex DNA  
10 may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it can contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of  
15 standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have recently been shown to be effective at  
20 inhibiting translation of mRNAs as well. (Wagner, R. (1994) *Nature* 372, 333). Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-coding regions of a *RAFTK* gene could be used in an antisense approach to inhibit translation of endogenous *RAFTK* mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start  
25 codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5', 3' or coding region of *RAFTK* mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In preferred  
30 embodiments, the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides, or at least 50 nucleotides. Oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the *RAFTK* nucleotide sequence of interest, are preferred.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives  
35 or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide

may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci.* 84, 648-652; PCT Publication No. WO 88/09810, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No. WO 89/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al. (1988) *BioTechniques* 6, 958-976) or intercalating agents. (See, e.g., Zon (1988) *Pharm. Res.* 5, 539-549). To this end, the oligonucleotide can be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide can comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide can also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier et al. (1987) *Nucl. Acids Res.* 15, 6625-6641). The

oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al. (1987) *Nucl. Acids Res.* 15, 6131-6148), or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215, 327-330).

Oligonucleotides of the invention can be synthesized by standard methods  
5 known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (1988, *Nucl. Acids Res.* 16, 3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al. (1988) *Proc.*  
10 *Natl. Acad. Sci. U.S.A.* 85, 7448-7451), etc.

While antisense nucleotides complementary to the *RAFTK* coding region sequence can be used, those complementary to the transcribed untranslated region are most preferred.

The antisense molecules are delivered to cells which express the *RAFTK in vivo*. A number of methods described herein and known in the art can be used for  
15 delivering the subject nucleic acids into to cells. A preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient results in the transcription of sufficient amounts of  
20 single stranded RNAs that forms complementary base pairs with the endogenous *RAFTK* transcripts and thereby prevent translation of the *RAFTK* mRNA. For example, a vector can be introduced *in vivo* such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the  
25 desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or  
30 constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon (1981) *Nature* 290, 304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al. (1980) *Cell* 22, 787-797), the herpes thymidine kinase promoter (Wagner et al. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78, 1441-1445), the regulatory sequences of the metallothionein  
35 gene (Brinster et al. (1982) *Nature* 296, 39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced directly into the tissue site; e.g., the choroid plexus or hypothalamus.

Alternatively, viral vectors can be used which selectively infect the desired tissue; (e.g., for brain, herpesvirus vectors may be used), in which case administration may be accomplished by another route (e.g., systematically).

Ribozyme molecules designed to catalytically cleave *RAFTK* mRNA transcripts can also be used to prevent translation of *RAFTK* mRNA and expression of *RAFTK*. (See, e.g., PCT International Publication WO90/11364, published October 4, 1990; Sarver et al. (1990) *Science* 247, 1222-1225). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy *RAFTK* mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988 *Nature*, 334, 585-591. There are numerous potential hammerhead ribozyme cleavage sites within the nucleotide sequence of human *RAFTK* cDNA. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the *RAFTK* mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena Thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al. (1984) *Science*, 224, 574-578; Zaug and Cech (1986) *Science*, 231, 470-475; Zaug, et al. (1986) *Nature*, 324, 429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech (1986) *Cell*, 47, 207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in *RAFTK*.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and are delivered to cells which express the *RAFTK* in vivo e.g., T cells. A preferred method of delivery involves using a DNA construct "encoding" the robozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells produce sufficient quantities of the ribozyme to destroy endogenous *RAFTK*

and inhibit translation. Because ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Endogenous *RAFTK* gene expression can also be reduced by inactivating or "knocking out" the *RAFTK* gene or its promoter using targeted homologous recombination. (E.g., see Smithies et al. (1985) *Nature* 317, 230-234; Thomas & Capecchi (1987) *Cell* 51, 503-512; Thompson et al. (1989) *Cell* 5, 313-321; each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional *RAFTK* (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous *RAFTK* gene (either the coding regions or regulatory regions of the *RAFTK* gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express *RAFTK in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the *RAFTK* gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive *RAFTK* (e.g., see Thomas & Capecchi (1987) and Thompson (1989), supra). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors, e.g., herpes virus vectors for delivery to brain tissue; e.g., the hypothalamus and/or choroid plexus.

Alternatively, endogenous *RAFTK* gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the *RAFTK* gene (i.e., the *RAFTK* promoter and/or enhancers) to form triple helical structures that prevent transcription of the *RAFTK* gene in target cells in the body. (See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6), 569-84; Helene, C., et al. (1992), *Ann. N.Y. Acad. Sci.* 660, 27-36; and Maher, L. J. (1992) *Bioassays* 14(12), 807-15).

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription are preferably single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides promotes triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules can be chosen that are purine-rich, for example, containing a stretch of g residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the



majority of the purine residues are located on a single strand of the targeted duplex, resulting in CGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizable stretch of either purines or pyrimidines to be present on one strand of a duplex.

Antisense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

The antisense constructs of the present invention, by antagonizing the normal biological activity of one of the *RAFTK* proteins, can be used in the manipulation of tissue survival, growth, migration, or differentiation, both *in vivo* and *ex vivo*. Furthermore, the anti-sense techniques (e.g. microinjection of antisense molecules, or transfection with plasmids whose transcripts are anti-sense with regard to a *RAFTK* mRNA or gene sequence) can be used to investigate role of *RAFTK* in developmental events, as well as the normal cellular function of *RAFTK* in adult tissue. Such techniques can be utilized in cell culture, but can also be used in the creation of transgenic animals, as detailed below.

*Polypeptides of the Present Invention*

The present invention also makes available isolated *RAFTK* polypeptides which are isolated from, or otherwise substantially free of other cellular proteins, especially other signal transduction factors and/or transcription factors which may normally be associated with the *RAFTK* polypeptide. The term "substantially free of other cellular proteins" (also referred to herein as "contaminating proteins") or "substantially pure or purified preparations" are defined as encompassing preparations of *RAFTK* polypeptides having less than about 20% (by dry weight) contaminating protein, and preferably having less than about 5% contaminating protein. Functional forms of the subject polypeptides can be prepared, for the first time, as purified preparations by using a cloned gene as described herein. The term "purified" when referring to a polypeptide or nucleic acid means that the polypeptide or nucleic acid is present in the substantial absence of other biological macromolecules, such as other proteins. The term "purified" as used herein preferably means at least 80% by dry weight, more preferably in the range of 95-99% by weight, and most preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above. "Isolated" and "purified" do not encompass either natural materials in their native state or natural materials that have been separated into components (e.g., in an acrylamide gel) but not obtained either as pure (e.g. lacking contaminating proteins, or chromatography reagents such as denaturing agents and polymers, e.g. acrylamide or agarose) substances or solutions. In preferred embodiments, purified *RAFTK* preparations lack any contaminating proteins from the same animal from which *RAFTK* is normally produced, as can be accomplished by recombinant expression of, for example, a human *RAFTK* protein in a non-human cell.

Full length proteins or fragments corresponding to one or more particular motifs and/or domains or to arbitrary sizes, for example, at least 5, 10, 25, 50, 75, 100, 125, 150 amino acids in length are within the scope of the present invention.

For example, isolated *RAFTK* polypeptides can include all or a portion of an amino acid sequences corresponding to a *RAFTK* polypeptide represented in one or more of one of SEQ ID NOs:2 or 4 and 4. Isolated peptidyl portions of *RAFTK* proteins can be obtained by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such peptides. In addition, fragments can be chemically synthesized using techniques known in the art such as

conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, a *RAFTK* polypeptide of the present invention can be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced  
5 (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of a wild-type (e.g., "authentic") *RAFTK* protein.

Another aspect of the present invention concerns recombinant forms of the *RAFTK* proteins. Recombinant polypeptides preferred by the present invention, in  
10 addition to native *RAFTK* proteins, are at least 60% homologous, preferably at least 70% and more preferably at least 80% homologous and even more preferably at least 90 % homologous with an amino acid sequence represented by one of SEQ ID NOs: 2 or 4. Polypeptides which are at least about 95% homologous with a sequence selected from the group consisting of SEQ ID NOs: 2 and 4 are also within the scope of the  
15 invention. In a preferred embodiment, a *RAFTK* protein of the present invention is a mammalian *RAFTK* protein. In a particularly preferred embodiment a *RAFTK* protein comprises the coding sequence of one of SEQ ID NOs:2 or 4. In particularly preferred embodiments, a *RAFTK* protein has a *RAFTK* bioactivity.

In certain preferred embodiments, the invention features a purified or  
20 recombinant *RAFTK* polypeptide having a molecular weight of approximately 115-125kD. In a preferred embodiment, the subject *RAFTK* polypeptide has a molecular weight of 123 kD. It will be understood that certain post-translational modifications, such as phosphorylation, can increase the apparent molecular weight of the *RAFTK* protein relative to the unmodified polypeptide chain.

The present invention further pertains to recombinant forms of one of the  
25 subject *RAFTK* polypeptides which are encoded by genes derived from a mammalian organism, and which have amino acid sequences evolutionarily related to the *RAFTK* proteins represented in one of SEQ ID NOs:2 or 4. Such recombinant *RAFTK* polypeptides preferably are capable of functioning in one of either role of an agonist  
30 or antagonist of at least one biological activity of a wild-type ("authentic") *RAFTK* protein of the appended sequence listing. The term "evolutionarily related to", with respect to amino acid sequences of mammalian *RAFTK* proteins, refers to both polypeptides having amino acid sequences which have arisen naturally, and also to mutational variants of mammalian *RAFTK* polypeptides which are derived, for  
35 example, by combinatorial mutagenesis. Such evolutionarily derived *RAFTK* polypeptides preferred by the present invention have a *RAFTK* bioactivity and are at least 60% homologous, preferably at least 70% homologous, and more preferably at

least 80% homologous and even more preferably at least 90% homologous with the amino acid sequence shown in one of SEQ ID NOs:2 or 4. Polypeptides at least 95-98% homologous are also within the scope of the invention. In a particularly preferred embodiment, a *RAFTK* protein comprises the amino acid coding sequence of one of SEQ ID NOs:2 or 4.

In general, polypeptides referred to herein as having a bioactivity of a mammalian *RAFTK* protein are defined as polypeptides which include an amino acid sequence corresponding (e.g., identical or homologous) to all or a portion of the amino acid sequences of a mammalian *RAFTK* proteins shown in one of SEQ ID NOs:2 or 4 and which mimic or antagonize all or a portion of the biological/biochemical activities of a naturally occurring *RAFTK* protein. Other biological activities of the subject *RAFTK* proteins are described herein or will be reasonably apparent to those skilled in the art. According to the present invention, a polypeptide has biological activity if it is a specific agonist or antagonist of a naturally-occurring form of a mammalian *RAFTK* protein.

The present invention further pertains to methods of producing the subject *RAFTK* polypeptides. For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding the subject polypeptides can be cultured under appropriate conditions to allow expression of the peptide to occur. The cells may be harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. The recombinant *RAFTK* polypeptide can be isolated from cell culture medium, host cells, or both, using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such peptide. In a preferred embodiment, the recombinant *RAFTK* polypeptide is a fusion protein containing a domain which facilitates its purification, such as GST fusion protein or poly(His) fusion protein.

Moreover, it will be generally appreciated that, under certain circumstances, it is advantageous to provide homologs of one of the subject *RAFTK* polypeptides which function in a limited capacity as one of either a *RAFTK* agonist (mimetic) or a *RAFTK* antagonist, in order to promote or inhibit only a subset of the biological activities of the naturally-occurring form of the protein. Thus, specific biological effects can be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of naturally occurring forms of *RAFTK* proteins.

Homologs of each of the subject *RAFTK* proteins can be generated by mutagenesis, such as by discrete point mutation(s), or by truncation. For example, mutation can give rise to homologs which retain substantially the same, or merely a subset, of the biological activity of the *RAFTK* polypeptide from which it was derived.

5 Alternatively, antagonistic forms of the protein can be generated which are able to inhibit the function of the naturally occurring form of the protein, such as by competitively binding to a downstream or upstream member of the *RAFTK* cascade which includes the *RAFTK* protein. In addition, agonistic forms of the protein may be generated which are constitutively active. Thus, the mammalian *RAFTK* protein and

10 homologs thereof provided by the subject invention may be either positive or negative regulators of *RAFTK* activity.

The recombinant *RAFTK* polypeptides of the present invention also include homologs of the authentic *RAFTK* proteins, such as versions of those protein which are resistant to proteolytic cleavage, as for example, due to mutations which alter

15 ubiquitination or other enzymatic targeting associated with the protein.

*RAFTK* polypeptides may also be chemically modified to create *RAFTK* derivatives by forming covalent or aggregate conjugates with other chemical moieties, such as glycosyl groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives of *RAFTK* proteins can be prepared by linking the chemical moieties to

20 functional groups on amino acid sidechains of the protein or at the N-terminus or at the C-terminus of the polypeptide.

Modification of the structure of the subject mammalian *RAFTK* polypeptides can be for such purposes as enhancing therapeutic or prophylactic efficacy, stability (e.g., ex vivo shelf life and resistance to proteolytic degradation in vivo), or post-translational modifications (e.g., to alter phosphorylation pattern of protein). Such

25 modified peptides, when designed to retain at least one activity of the naturally-occurring form of the protein, or to produce specific antagonists thereof, are considered functional equivalents of the *RAFTK* polypeptides described in more detail herein. Such modified peptides can be produced, for instance, by amino acid

30 substitution, deletion, or addition.

For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e. isosteric and/or isoelectric mutations) will not have a major effect on the

35 biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids are can be divided into four families: (1) acidic =

- aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar = alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. In similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine histidine, (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tryptophan; (5) amide = asparagine, glutamine; and (6) sulfur -containing = cysteine and methionine. (sec. for example, Biochemistry, 2nd ed., Ed. by L. Stryer, WH Freeman and Co.: 1981). Whether a change in the amino acid sequence of a peptide results in a functional *RAFTK* homolog (e.g. functional in the sense that the resulting polypeptide mimics or antagonizes the wild-type form) can be readily determined by assessing the ability of the variant peptide to produce a response in cells in a fashion similar to the wild-type protein, or competitively inhibit such a response.
- Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

This invention further provides a method for generating sets of combinatorial mutants of the subject *RAFTK* proteins as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g. homologs) that modulate a *RAFTK* bioactivity. The purpose of screening such combinatorial libraries is to generate, for example, novel *RAFTK* homologs which can act as either agonists or antagonist, or alternatively, possess novel activities all together. To illustrate, combinatorially-derived homologs can be generated to have an increased potency relative to a naturally occurring form of the protein.

Likewise, *RAFTK* homologs can be generated by the present combinatorial approach to selectively inhibit (antagonize) an authentic *RAFTK*. For instance, mutagenesis can provide *RAFTK* homologs which are able to bind other proteins in a *RAFTK* signaling pathway yet prevent propagation of the signal, e.g. the homologs can be dominant negative mutants. Moreover, manipulation of certain domains of *RAFTK* by the present method can provide domains more suitable for use in fusion proteins.

In one embodiment, the variegated library of *RAFTK* variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential *RAFTK* sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display) containing the set of *RAFTK* sequences therein.

There are many ways by which such libraries of potential *RAFTK* homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential *RAFTK* sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, S. A. (1983) *Tetrahedron* 39, 3; Itakura et al. (1981) *Recombinant DNA, Proc 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp 273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53, 323. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) *Science* 249, 386-390; Roberts et al. (1992) *PNAS* 89, 2429-2433; Devlin et al. (1990) *Science* 249, 404-406).

Likewise, a library of coding sequence fragments can be provided for a *RAFTK* clone in order to generate a variegated population of *RAFTK* fragments for screening and subsequent selection of bioactive fragments. A variety of techniques are known in the art for generating such libraries, including chemical synthesis. In one embodiment, a library of coding sequence fragments can be generated by (i) treating a double stranded PCR fragment of a *RAFTK* coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule; (ii) denaturing the double stranded DNA; (iii) renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products; (iv) removing single stranded portions from reformed duplexes by treatment with S1 nuclease; and (v) ligating the resulting fragment library into an expression vector. By this exemplary method, an expression library can be derived which codes for N-terminal, C-terminal and internal fragments of various sizes.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a certain property. Such techniques are generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of *RAFTK* homologs. The most widely used techniques for screening large gene libraries typically comprise cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays

described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate *RAFTK* sequences created by combinatorial mutagenesis techniques.

5 In one embodiment, cell based assays can be exploited to analyze the variegated *RAFTK* library. For instance, the library of expression vectors can be transfected into a cell line ordinarily responsive to a ligand which transduces signals via a pathway involving *RAFTK*, such as, for example, stem cell factor, thrombin, fibronectin, CSF-1/M-CSF, T cell receptor stimulation, bFGF, oncoprotein M, IL-6, or TNF $\alpha$ . The transfected cells are then contacted with a ligand and the effect of the *RAFTK* mutant can be detected, e.g. on cell viability. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of a *RAFTK* activity, and the individual clones further characterized.

15 Combinatorial mutagenesis has a potential to generate very large libraries of mutant proteins, e.g., in the order of  $10^{26}$  molecules. Combinatorial libraries of this size may be technically challenging to screen even with high throughput screening assays. To overcome this problem, a new technique has been developed recently, recursive ensemble mutagenesis (REM), which allows one to avoid the very high proportion of non-functional proteins in a random library and simply enhances the frequency of functional proteins, thus decreasing the complexity required to achieve a useful sampling of sequence space. REM is an algorithm which enhances the frequency of functional mutants in a library when an appropriate selection or screening method is employed (Arkin and Yourvan, 1992, *PNAS USA* 89, 7811-7815; Yourvan et al. (1992) *Parallel Problem Solving from Nature*, 2, In Maenner and Manderick, eds., Elsevir Publishing Co., Amsterdam, pp. 401-410; Delgrave et al. 20 (1993) *Protein Engineering* 6(3), 327-331).

The invention also provides for reduction of the mammalian *RAFTK* proteins to generate mimetics, e.g. peptide or non-peptide agents, which are able to disrupt binding of a mammalian *RAFTK* polypeptide of the present invention with either upstream or downstream components of a TGF $\beta$  signaling cascade, such as binding 30 proteins or interactors. Thus, such mutagenic techniques as described above are also useful to map the determinants of the *RAFTK* proteins which participate in protein-protein interactions involved in, for example, binding of the subject mammalian *RAFTK* polypeptide to proteins which may function upstream (including both activators and repressors of its activity) or to proteins or nucleic acids which may 35 function downstream of the *RAFTK* polypeptide, whether they are positively or negatively regulated by it. To illustrate, the critical residues of a subject *RAFTK* polypeptide which are involved in molecular recognition of binding proteins upstream



or downstream of a *RAFTK* can be determined and used to generate *RAFTK*-derived peptidomimetics which competitively inhibit binding of the authentic *RAFTK* protein with that moiety. By employing, for example, scanning mutagenesis to map the amino acid residues of each of the subject *RAFTK* proteins which are involved in binding other extracellular proteins, peptidomimetic compounds can be generated which mimic those residues of the *RAFTK* protein which facilitate the interaction. Such mimetics may then be used to interfere with the normal function of a *RAFTK* protein. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29, 295; and Ewenson et al. in *Peptides: Structure and Function (Proceedings of the 9th American Peptide Symposium)* Pierce Chemical Co. Rockland, IL, 1985), b-turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26, 647; and Sato et al. (1986) *J Chem Soc Perkin Trans* 1, 1231), and b-aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun* 126, 419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134, 71).

#### *Cells expressing recombinant RAFTK polypeptides.*

This invention also pertains to a host cell transfected to express a recombinant form of the subject *RAFTK* polypeptides. The host cell can be any prokaryotic or eukaryotic cell. Thus, a nucleotide sequence derived from the cloning of mammalian *RAFTK* proteins, encoding all or a selected portion of the full-length protein, can be used to produce a recombinant form of a mammalian *RAFTK* polypeptide via microbial or eukaryotic cellular processes. Ligating the polynucleotide sequence into a gene construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), are standard procedures used in producing other well-known proteins, e.g. MAP kinase, p53, WT1, PTP phosphatases, SRC, and the like. Similar procedures, or modifications thereof, can be employed to prepare recombinant *RAFTK* polypeptides by microbial means or tissue-culture technology in accord with the subject invention.

The recombinant *RAFTK* genes can be produced by ligating nucleic acid encoding a *RAFTK* protein, or a portion thereof, into a vector suitable for expression in either prokaryotic cells, eukaryotic cells, or both. Expression vectors for

production of recombinant forms of the subject *RAFTK* polypeptides include plasmids and other vectors. For instance, suitable vectors for the expression of a *RAFTK* polypeptide include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al. (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83, incorporated by reference herein). These vectors can replicate in *E. coli* due the presence of the pBR322 ori. and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used. In an illustrative embodiment, a *RAFTK* polypeptide is produced recombinantly utilizing an expression vector generated by sub-cloning the coding sequence of one of the *RAFTK* genes represented in one of SEQ ID NOs:1 or 3.

The preferred mammalian expression vectors contain both prokaryotic sequences, to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNA1/amp, pcDNA1/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papillomavirus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989) Chapters 16 and 17.

In some instances, it is desirable to express the recombinant *RAFTK* polypeptide by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the  $\beta$ -gal containing pBlueBac III).

*Fusion proteins and Immunogens.*

In another embodiment, the coding sequences for the polypeptide can be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. This type of expression system can be useful under conditions where it is desirable to produce an immunogenic fragment of a *RAFTK* protein. For example, the VP6 capsid protein of rotavirus can be used as an immunologic carrier protein for portions of the *RAFTK* polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of a subject *RAFTK* protein to which antibodies are to be raised can be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising *RAFTK* epitopes as part of the virion. It has been demonstrated with the use of immunogenic fusion proteins utilizing the Hepatitis B surface antigen fusion proteins that recombinant Hepatitis B virions can be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of a *RAFTK* protein and the poliovirus capsid protein can be created to enhance immunogenicity of the set of polypeptide antigens (see, for example, EP Publication No: 0259149; and Evans et al. (1989) *Nature* 339, 385; Huang et al. (1988) *J. Virol.* 62, 3855; and Schlienger et al. (1992) *J. Virol.* 66, 2).

The Multiple Antigen Peptide system for peptide-based immunization can also be utilized to generate an immunogen, wherein a desired portion of a *RAFTK* polypeptide is obtained directly from organo-chemical synthesis of the peptide onto an oligomeric branching lysine core (see, for example, Posnett et al. (1988) *JBC* 263, 1719 and Nardelli et al. (1992) *J. Immunol.* 148, 914). Antigenic determinants of *RAFTK* proteins can also be expressed and presented by bacterial cells.

In addition to utilizing fusion proteins to enhance immunogenicity, it is widely appreciated that fusion proteins can also facilitate the expression of proteins, and accordingly, can be used in the expression of the mammalian *RAFTK* polypeptides of the present invention. For example, *RAFTK* polypeptides can be generated as glutathione-S-transferase (GST-fusion) proteins. Such GST-fusion proteins can enable easy purification of the *RAFTK* polypeptide, as for example by the use of glutathione-derivatized matrices (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. (N.Y.: John Wiley & Sons, 1991)).

In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, can allow purification of the

expressed fusion protein by affinity chromatography using a Ni<sup>2+</sup> metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase to provide the purified protein (e.g., see Hochuli et al. (1987) *J. Chromatography* 411, 177; and Janknecht et al. *PNAS* 88, 8972).

- 5           Techniques for making fusion genes are known to those skilled in the art. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline  
10   phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs  
15   between two consecutive gene fragments which can subsequently be annealed to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. John Wiley & Sons: 1992).

#### *Antibodies*

- 20           Another aspect of the invention pertains to antibodies specifically reactive with a vertebrate *RAFTK* protein, preferably a mammalian *RAFTK* protein. For example, by using immunogens derived from a *RAFTK* protein, e.g. based on the cDNA sequences, anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, *Antibodies: A Laboratory Manual* ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a  
25   mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide (e.g., a mammalian *RAFTK* polypeptide or an antigenic fragment which is capable of eliciting an antibody response, or a fusion protein as described above). Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic  
30   portion of a *RAFTK* protein can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of a *RAFTK* protein of a  
35   mammal, e.g. antigenic determinants of a protein represented by one of SEQ ID NOs:2 or 4.

Following immunization of an animal with an antigenic preparation of a *RAFTK* polypeptide, anti- *RAFTK* antisera can be obtained and, if desired, polyclonal anti- *RAFTK* antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) *Nature*, 256, 495-497), the human B cell hybridoma technique (Kozbar et al., (1983) *Immunology Today*, 4, 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a mammalian *RAFTK* polypeptide of the present invention and monoclonal antibodies isolated from a culture comprising such hybridoma cells.

The term "antibody" as used herein is intended to include fragments thereof which are also specifically reactive with one of the subject mammalian *RAFTK* polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab)<sub>2</sub> fragments can be generated by treating antibody with pepsin. The resulting F(ab)<sub>2</sub> fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the present invention is further intended to include bispecific and chimeric molecules having affinity for a *RAFTK* protein conferred by at least one CDR region of the antibody.

Antibodies which specifically bind *RAFTK* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of each of the subject *RAFTK* polypeptides. Anti-*RAFTK* antibodies can be used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *RAFTK* protein levels in tissue as part of a clinical testing procedure. Likewise, the ability to monitor *RAFTK* protein levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. Diagnostic assays using anti- *RAFTK* antibodies can include, for example, immunoassays designed to aid in early diagnosis of a degenerative disorder, particularly ones which are manifest at birth. Diagnostic assays using anti- *RAFTK* polypeptide antibodies can also include immunoassays designed to aid in early diagnosis and phenotyping neoplastic or hyperplastic disorders.

Another application of anti-*RAFTK* antibodies of the present invention is in the immunological screening of cDNA libraries constructed in expression vectors

such as  $\lambda$ gt11,  $\lambda$ gt18-23,  $\lambda$ ZAP, and  $\lambda$ ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance,  $\lambda$ gt11 will produce fusion proteins whose amino termini consist of  $\beta$ -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a *RAFTK* protein, e.g. other orthologs of a particular *RAFTK* protein or other paralogs from the same species, can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-*RAFTK* antibodies. Positive phage detected by this assay can then be isolated from the infected plate. Thus, the presence of *RAFTK* homologs can be detected and cloned from other animals, as can alternate isoforms (including splicing variants) from humans.

#### *Methods of Treating Disease*

There are a wide variety of pathological conditions which can be treated using the *RAFTK* therapeutics of the present invention. For example, *RAFTK* therapeutics which modulate *RAFTK* activity in B cells, T cells, and monocytes can be used to treat immune-mediated disorders and mediate both cell mediated and humoral immune responses.

Normal hematopoietic cells are dependent on growth factors for growth and differentiation and the loss of this growth factor dependence can lead to autonomous growth. The involvement of *RAFTK* in several growth factor signaling pathways indicates that missexpression of *RAFTK* can lead to the development of cancers, and the present invention contemplates modulating *RAFTK* expression and/or activity to control aberrant cell growth. In a preferred embodiment *RAFTK* is modulated to treat cancers of hematopoietic cells. In another embodiment malignancy can be suppressed in certain cells e.g., leukemic cells, by modulating *RAFTK* to induce cellular differentiation in "differentiation therapy", for example, in the treatment of leukemia, as has been demonstrated with cytokines or other compounds (Sachs (1996) *Proc. Natl. Acad. Sci. USA* 93:4742).

The subject *RAFTK* proteins can also be modulated to either induce or inhibit apoptosis in a cell. In certain embodiments the subject *RAFTK* proteins can be manipulated to induce apoptosis in cancer cells. In some embodiments *RAFTK* can be modulated in a patient in conjunction with other cancer therapies. Alternatively, in instances when it is desirable to inhibit apoptosis, such as apoptosis induced by chemotherapeutic compounds and irradiation, *RAFTK* may be modulated to inhibit apoptosis.

Cytoskeletal rearrangement has been correlated with growth control and gene expression and is critical in cell adhesion and migration and the modulation of *RAFTK* bioactivity can alter cellular functions which depend upon the cytoskeleton, including, for example, normal tissue maintenance and proliferation and tissue remodeling which occur in response to injury (Turner et al. (1995) *J. Cell Science* 108:333). In a preferred embodiment, *RAFTK* bioactivity is modulated to reduce metastasis of a cancer cell.

Yet another aspect of the present invention pertains to the therapeutic application of a *RAFTK* therapeutic to enhance survival of neurons and other neuronal cells in both the central nervous system and the peripheral nervous system. The expression of *RAFTK* in neuronal cells and their role in signaling pathways involved in apoptosis (Tokiwa et al. (1996) *Science* 273:792) indicates that certain of the *RAFTK* proteins participate in control of adult neurons with regard to maintenance, functional performance, and aging of normal cells; repair and regeneration processes in chemically or mechanically lesioned cells; and prevention of degeneration and premature death which result from loss of differentiation in certain pathological conditions. The present invention specifically contemplates applications of the subject method to the treatment of (prevention and/or reduction of the severity of) neurological conditions deriving from: (i) acute, subacute, or chronic injury to the nervous system, including traumatic injury, chemical injury, vascular injury and deficits (such as the ischemia resulting from stroke), together with infectious/ inflammatory and tumor-induced injury; (ii) aging of the nervous system including Alzheimer's disease; (iii) chronic neurodegenerative diseases of the nervous system, including Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis and the like, as well as spinocerebellar degenerations; and (iv) chronic immunological diseases of the nervous system or affecting the nervous system, including multiple sclerosis.

Many neurological disorders are associated with degeneration of discrete populations of neuronal elements and can be treatable with a therapeutic regimen which includes a *RAFTK* therapeutic. For example, Alzheimer's disease is associated with deficits in several neurotransmitter systems, both those that project to the neocortex and those that reside with the cortex. For instance, the nucleus basalis in patients with Alzheimer's disease have been observed to have a profound (75%) loss of neurons compared to age-matched controls. Although Alzheimer's disease is by far the most common form of dementia, several other disorders can produce dementia. Several of these are degenerative diseases characterized by the death of neurons in various parts of the central nervous system, especially the cerebral cortex. However, some forms of dementia are associated with degeneration of the thalamus or the white

matter underlying the cerebral cortex. Here, the cognitive dysfunction results from the isolation of cortical areas by the degeneration of efferents and afferents.

Huntington's disease involves the degeneration of intrastriatal and cortical cholinergic neurons and GABAergic neurons. Pick's disease is a severe neuronal degeneration in

5 the neocortex of the frontal and anterior temporal lobes, sometimes accompanied by death of neurons in the striatum. Treatment of patients suffering from such degenerative conditions can include the application of *RAFTK* therapeutics, in order to control, for example, differentiation and apoptotic events which give rise to loss of neurons (e.g. to enhance survival of existing neurons) as well as promote

10 differentiation and repopulation by progenitor cells in the area affected.

In addition to degenerative-induced dementias, a pharmaceutical preparation of one or more of the subject *RAFTK* therapeutics can be applied opportunely in the treatment of neurodegenerative disorders which have manifestations of tremors and involuntary movements. Parkinson's disease, for example, primarily affects

15 subcortical structures and is characterized by degeneration of the nigrostriatal pathway, raphe nuclei, locus cereleus, and the motor nucleus of vagus. Ballism is typically associated with damage to the subthlamic nucleus, often due to acute vascular accident.

Also included in the methods of the invention are treatment of neurogenic and

20 myopathic diseases which ultimately affect the somatic division of the peripheral nervous system and are manifest as neuromuscular disorders. In an illustrative embodiment, the subject method is used to treat amyotrophic lateral sclerosis. ALS is a name given to a complex of disorders that comprise upper and lower motor neurons. Patients may present with progressive spinal muscular atrophy, progressive

25 bulbar palsy, primary lateral sclerosis, or a combination of these conditions. The major pathological abnormality is characterized by a selective and progressive degeneration of the lower motor neurons in the spinal cord and the upper motor neurons in the cerebral cortex. The therapeutic application of a *RAFTK* therapeutic, can be used alone, or in conjunction with neurotrophic factors such as CNTF, BDNF

30 or NGF to prevent and/or reverse motor neuron degeneration in ALS patients.

Another aspect of the present invention relates to a method of inducing and/or maintaining a differentiated state or enhancing survival by contacting the cells with an agent which modulates *RAFTK*-dependent signaling by a growth factor. For instance, it is contemplated by the invention that, in light of the present finding of a broad

35 involvement of *RAFTK* proteins in signal transduction in a variety of different cell types, the subject *RAFTK* signaling molecules can be used in a wide range of therapeutic regimens both *in vitro* and *in vivo*. A "*RAFTK* therapeutic" can be, as



appropriate, any of the preparations described above, including isolated polypeptides, gene therapy constructs, antisense molecules, peptidomimetics or agents identified in the drug assays provided herein.

5 In one embodiment the *RAFTK* proteins of the present invention can modulate the differentiation or maturation of hematopoietic cells; the subject *RAFTK* polypeptides are capable of influencing both the differentiation and maturation of pluripotent stem cells and the proliferation of differentiated cells. In a preferred embodiment *RAFTK* bioactivity is modulated in CD34+ bone marrow cells; the presence of the cell-surface marker CD34 in humans has been found to correlate with  
10 bone marrow progenitors which proliferate to hematopoietic cytokines.

Hematopoiesis can be modulated either *in vitro* or *in vivo* and the subject *RAFTK* therapeutics can be used alone or in combination with cytokines and/or colony stimulating factors. For example, in certain embodiments, it may be desirable to coadminister a growth factor, for example, G-CSF and/or IL-3 (Lemoli et al. Experimental Hematology 1995, 23:1520) or SCF which has been shown to act with  
15 other cytokines to stimulate hematopoietic colony formation Martin et al. (1990) *Cell* 63, 203), stimulate hematopoiesis (Andrews et al. (1991) *Blood* 78, 1975), and rescue from the effects of lethal irradiation (Zsebo et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 9464) can be combined with with the subject *RAFTK* therapeutics. When stem  
20 cells are induced to mature and/or proliferate *in vitro*, the subject *RAFTK* therapeutics can be combined with culture of the stem cells on feeder cells. Stem cells in which *RAFTK* is modulated can be useful, for example in the expansion of cells for autologous or allogeneic transplantation of stem cells or differentiated cells. *RAFTK* can be modulated to enhance engraftment and/or hematopoiesis after allogeneic bone  
25 marrow transplantation.

In preferred embodiments *RAFTK* can be modulated to control megakaryocyte development and to ameliorate diseases caused by abnormalities in megakaryocytic cells, for example, thrombocytopenia, myelodysplastic syndrome, myeloproliferative disorder, aplastic anemia, chronic myelogenous leukemia. Platelets are derived from  
30 megakaryocytes, and the subject *RAFTK* molecules can be used to correct abnormalities in platelet number or function. Platelets are important in numerous hemorrhagic and thrombotic disorders. In a preferred embodiment *RAFTK* bioactivity can be modulated to control platelet aggregation. In yet another embodiment, the subject *RAFTK* therapeutics can play a role in the maturation of cells of the erythroid  
35 lineage.

Among the approaches which can be used to ameliorate disease symptoms involving the misexpression of a *RAFTK* gene are, for example, antisense, ribozyme.

and triple helix molecules described above. Compounds that compete with an *RAFTK* protein for binding with an active portion of *RAFTK* will antagonize a *RAFTK* protein, thereby inducing a therapeutic effect. Examples of suitable compounds include the antagonists or homologues described in detail above. In other instances, the increased expression or bioactivity of a *RAFTK* protein may be desirable and may be accomplished by, for example the use of the *RAFTK* agonists or mimetics or by gene replacement therapy, as described herein.

#### *Effective Dose*

It is within the level of ordinary skill in the art to determine dosages of the subject *RAFTK* therapeutics. The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED<sub>50</sub> with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC<sub>50</sub> (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans.

#### *Formulation and Use*

Pharmaceutical compositions for use in accordance with the present invention can be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds (e.g., *RAFTK* polypeptides or *RAFTK* nucleic acids) and their physiologically acceptable salts and solvates may be formulated for administration by, for example, injection, inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For such therapy, the oligomers of the invention can be formulated for a variety of loads of administration, including systemic and topical or localized administration. Techniques and formulations generally can be found in *Remington's Pharmaceutical Sciences*, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous.

For oral administration, the pharmaceutical compositions can take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art.

Preparations for oral administration may be suitably formulated to give controlled release of the active compound.

The compounds can be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions can take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For topical administration, the oligomers of the invention are formulated into ointments, salves, gels, or creams as generally known in the art.

In clinical settings, the gene delivery systems for the therapeutic *RAFTK* gene can be introduced into a patient by any of a number of methods, each of which is familiar in the art. For example, a pharmaceutical preparation of the gene delivery system can be introduced systemically, e.g. by intravenous injection, and specific transduction of the protein in the target cells occurs predominantly from specificity of transfection provided by the gene delivery vehicle, cell-type or tissue-type expression due to the transcriptional regulatory sequences controlling expression of the receptor gene, or a combination thereof. In other embodiments, initial delivery of the recombinant gene is localized. For example, the gene delivery vehicle can be introduced by catheter (see U.S. Patent 5,328,470) or by stereotactic injection (e.g. Chen et al. (1994) *PNAS* 91, 3054-3057). A mammalian *RAFTK* gene, such as any

one of the sequences represented in one of SEQ ID NOS:1 or 3, or a sequence homologous thereto can be delivered in a gene therapy construct by electroporation using techniques described, for example, by Dev et al. ((1994) *Cancer Treat Rev* 20, 105-115).

5       The pharmaceutical preparation of the gene therapy construct can include the gene delivery system in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery system can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can comprise one or more cells which produce  
10       the gene delivery system.

      The compositions can, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device can be accompanied by instructions for administration.

15

#### *Diagnostic and Prognostic Assays*

      The present method provides a method for determining if a subject is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In preferred embodiments, the methods can be characterized as comprising detecting, in  
20       a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of (i) an alteration affecting the integrity of a gene encoding a *RAFTK*-protein, or (ii) the mis-expression of the *RAFTK* gene. To illustrate, such genetic lesions can be detected by ascertaining the existence of at least one of (i) a deletion of one or more nucleotides from a *RAFTK* gene, (ii) an addition  
25       of one or more nucleotides to a *RAFTK* gene, (iii) a substitution of one or more nucleotides of a *RAFTK* gene, (iv) a gross chromosomal rearrangement of a *RAFTK* gene, (v) a gross alteration in the level of a messenger RNA transcript of a *RAFTK* gene, (vi) aberrant modification of a *RAFTK* gene, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild type splicing pattern of a  
30       messenger RNA transcript of a *RAFTK* gene, (viii) a non-wild type level of a *RAFTK*-protein, (ix) allelic loss of a *RAFTK* gene, and (x) inappropriate post-translational modification of a *RAFTK*-protein. As described herein, the present invention provides a large number of assay techniques for detecting lesions in a *RAFTK* gene, and importantly, provides the ability to discern between different molecular causes  
35       underlying *RAFTK*-dependent aberrant cell growth, proliferation and/or differentiation.

In an exemplary embodiment, there is provided a nucleic acid composition comprising a (purified) oligonucleotide probe including a region of nucleotide sequence which is capable of hybridizing to a sense or antisense sequence of a *RAFTK* gene, such as represented by one of SEQ ID NOs:1 or 3, or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with the subject *RAFTK* genes or naturally occurring mutants thereof. The nucleic acid of a cell is rendered accessible for hybridization, the probe is exposed to nucleic acid of the sample, and the hybridization of the probe to the sample nucleic acid is detected. Such techniques can be used to detect lesions at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels.

In certain embodiments, detection of the lesion comprises utilizing the probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) *Science* 241, 1077-1080; and Nakazawa et al. (1994) *PNAS* 91, 360-364), the latter of which can be particularly useful for detecting point mutations in the *RAFTK*-gene (see Abravaya et al. (1995) *Nuc Acid Res* 23, 675-682). In an illustrative embodiment, the method includes the steps of (i) collecting a sample of cells from a patient, (ii) isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, (iii) contacting the nucleic acid sample with one or more primers which specifically hybridize to a *RAFTK* gene under conditions such that hybridization and amplification of the *RAFTK*-gene (if present) occurs, and (iv) detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein. Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. et al. (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874-1878), transcriptional amplification system (Kwoh, D.Y. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1173-1177), Q-Beta Replicase (Lizardi, P.M. et al. (1988) *Bio/Technology* 6, 1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In a preferred embodiment of the subject assay, mutations in a *RAFTK* gene from a sample cell are identified by alterations in restriction enzyme cleavage

patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the *RAFTK* gene and detect mutations by comparing the sequence of the sample *RAFTK* with the corresponding wild-type (control) sequence. Exemplary sequencing reactions include those based on techniques developed by Maxim and Gilbert (*Proc. Natl Acad Sci USA* (1977) 74, 560) or Sanger (Sanger et al (1977) *Proc. Nat. Acad. Sci* 74, 5463). It is also contemplated that any of a variety of automated sequencing procedures may be utilized when performing the subject assays ((1995) *Biotechniques* 19, 448), including by sequencing by mass spectrometry (see, for example PCT publication WO 94/16101; Cohen et al. (1996) *Adv Chromatogr* 36, 127-162; and Griffin et al. (1993) *Appl Biochem Biotechnol* 38, 147-159). It will be evident to one skilled in the art that, for certain embodiments, the occurrence of only one, two or three of the nucleic acid bases need be determined in the sequencing reaction. For instance, A-tract or the like, e.g., where only one nucleic acid is detected, can be carried out.

In a further embodiment, protection from cleavage agents (such as a nuclease, hydroxylamine or osmium tetroxide and with piperidine) can be used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers, et al. (1985) *Science* 230, 1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labelled) RNA or DNA containing the wild-type *RAFTK* sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton et al (1988) *Proc. Natl Acad Sci USA* 85, 4397; Saleeba et al (1992) *Methods Enzymol.* 217, 286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in *RAFTK* cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al. (1994) *Carcinogenesis* 15, 1657-1662). According to an exemplary embodiment, a probe based on a *RAFTK* sequence, e.g., a wild-type *RAFTK* sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility is used to identify mutations in *RAFTK* genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) *Proc Natl. Acad. Sci USA* 86, 2766, see also Cotton (1993) *Mutat Res* 285, 125-144; and Hayashi (1992) *Genet Anal Tech Appl* 9, 73-79). Single-stranded DNA fragments of sample and control *RAFTK* nucleic acids is denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labelled or detected with labelled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) *Trends Genet* 7, 5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al (1985) *Nature* 313, 495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing agent gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265, 12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or

selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al. (1986) *Nature* 324, 163); Saiki et al (1989) *Proc. Natl Acad. Sci USA* 86, 6230).

- 5 Such allele specific oligonucleotide hybridization techniques may be used to test one mutation per reaction when oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labelled target DNA.

- Alternatively, allele specific amplification technology which depends on  
10 selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al (1989) *Nucleic Acids Res.* 17, 2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent, or  
15 reduce polymerase extension (Prossner (1993) *Tibtech* 11, 238. In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al (1992) *Mol. Cell Probes* 6, 1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88, 189). In  
20 such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

- Another embodiment of the invention provides for a nucleic acid composition comprising a (purified) oligonucleotide probe including a region of nucleotide  
25 sequence which is capable of hybridizing to a sense or antisense sequence of a *RAFTK*-gene, or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with the subject *RAFTK*-genes or naturally occurring mutants thereof. The nucleic acid of a cell is rendered accessible for hybridization, the probe is exposed to nucleic acid of the sample, and the  
30 hybridization of the probe to the sample nucleic acid is detected. Such techniques can be used to detect lesions at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels. Such oligonucleotide probes can be used for both predictive and therapeutic evaluation of allelic mutations which might be manifest in, for example, neoplastic or hyperplastic  
35 disorders (e.g. aberrant cell growth).

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody



reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a *RAFTK* gene.

Diagnostic procedures may also be performed *in situ* directly upon tissue sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents may be used as probes and/or primers for such *in situ* procedures (see, for example, Nuovo, G.J., 1992, *PCR in situ hybridization: protocols and applications*, Raven Press, NY).

In yet another embodiment mutant *RAFTK* proteins can be detected using the protein truncation test (PTT) (Dowton and Salugh. 1995, *Clin. Chem* 41:785). For PTT, RNA is initially isolated and reverse-transcribed, and the segment of interest is amplified by PCR. The PCR products are then used as a template for nested PCR amplification with a primer containing an RNA polymerase promoter and a translation initiation sequence. After amplification, the unique motifs incorporated into the primer permit sequential *in vitro* transcription and translation of the PCR products. Protein products are analyzed by electrophoresis and mutantations which cause truncation of the protein are identified by a change in the molecular weight of the protein. DNA may also be used.

In addition to methods which focus primarily on the detection of one nucleic acid sequence, profiles may also be assessed in such detection schemes. Fingerprint profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR.

Antibodies directed against wild type or mutant *RAFTK* proteins, which are discussed, above, may also be used indisease diagnostics and prognostics. Such diagnostic methods, may be used to detect abnormalities in the level of *RAFTK* protein expression, or abnormalities in the structure and/or tissue, cellular, or subcellular location of *RAFTK* protein. Structural differences may include, for example, differences in the size, electronegativity, or antigenicity of the mutant *RAFTK* protein relative to the normal *RAFTK* protein. Protein from the tissue or cell type to be analyzed may easily be detected or isolated using techniques which are well known to one of skill in the art, including but not limited to western blot analysis. For a detailed explanation of methods for carrying out western blot analysis, see Sambrook et al, 1989, *supra*, at Chapter 18. The protein detection and isolation methods employed herein may also be such as those described in Harlow and Lane, for example, (Harlow, E. and Lane, D., 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), which is incorporated herein by reference in its entirety.

This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below) coupled with light microscopic, flow cytometric, or fluorimetric detection. The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for in situ detection of *RAFTK* proteins. In situ detection may be accomplished by removing a histological specimen from a patient, and contacting it with a labeled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the *RAFTK* protein, but also its distribution in the examined tissue. Using the present invention, one of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such in situ detection.

Often a solid phase support or carrier is used as a support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amyloses, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

One means for labeling an anti- *RAFTK* protein specific antibody is via linkage to an enzyme and use in an enzyme immunoassay (EIA) (Voller (1978), "The Enzyme Linked Immunosorbent Assay (ELISA)", *Diagnostic Horizons* 2, 1-7, 1978, Microbiological Associates Quarterly Publication, Walkersville, MD; Voller, et al., J. Clin. Pathol. 31, 507-520 (1978); Butler (1981) *Meth. Enzymol.* 73, 482-523; Maggio, (ed.) *Enzyme Immunoassay*, CRC Press, Boca Raton, FL, 1980; Ishikawa, et al., (eds.) *Enzyme Immunoassay*, Kaku Shoin, Tokyo, 1981). The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means.

Enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

10 Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect fingerprint gene wild type or mutant peptides through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., *Principles of Radioimmunoassays*, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986, which is incorporated by reference herein). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

20 It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine. The antibody can also be labeled using fluorescence emitting metals such as  $^{152}\text{Eu}$ , or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

30 The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, theromatic acridinium ester, imidazole, acridinium salt and oxalate ester. Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

Moreover, it will be understood that any of the above methods for detecting alterations in a *RAFTK* gene or gene product can be used to monitor the course of treatment or therapy.

## 5     *Drug Screening Assays*

Furthermore, by making available purified and recombinant *RAFTK* polypeptides, the present invention facilitates the development of assays which can be used to screen for compounds, including *RAFTK* homologs, which are either agonists or antagonists of the normal cellular function of the subject *RAFTK* polypeptides, or  
10     of their role in the pathogenesis of cellular differentiation and/or proliferation and disorders related thereto. A variety of assay formats can be utilized and, in light of the present inventions, will be comprehended by a skilled artisan.

### *Cell-free assays*

15     In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid  
20     development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the in vitro system, the assay instead can be focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with upstream or  
25     downstream elements. Accordingly, in an exemplary screening assay of the present invention, the compound of interest is contacted with proteins which may function upstream (including both activators and repressors of its activity) or to proteins or nucleic acids which may function downstream of the *RAFTK* polypeptide, whether they are positively or negatively regulated by it. To the mixture of the compound and  
30     the upstream or downstream element is then added a composition containing a *RAFTK* polypeptide. Detection and quantification of complexes of *RAFTK* with its upstream or downstream elements provide a means for determining a compound's efficacy at inhibiting (or potentiating) complex formation between *RAFTK* and the *RAFTK*-binding elements. The efficacy of the compound can be assessed by  
35     generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, isolated and purified *RAFTK*

polypeptide is added to a composition containing the *RAFTK*-binding element, and the formation of a complex is quantitated in the absence of the test compound.

Complex formation between the *RAFTK* polypeptide and a *RAFTK* binding element may be detected by a variety of techniques. Modulation of the formation of  
5 complexes can be quantitated using, for example, detectably labeled proteins such as radiolabeled, fluorescently labeled, or enzymatically labeled *RAFTK* polypeptides, by immunoassay, or by chromatographic detection.

Typically, it is desirable to immobilize either *RAFTK* or its binding protein to facilitate separation of complexes from uncomplexed forms of one or both of the  
10 proteins, as well as to accommodate automation of the assay. Binding of *RAFTK* to an upstream or downstream element, in the presence and absence of a candidate agent, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be  
15 bound to a matrix. For example, glutathione-S-transferase/*RAFTK* (*GST/RAFTK*) fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g. <sup>35</sup>S-labeled) and the test compound, and the mixture incubated under conditions conducive to complex formation, for example at  
20 physiological conditions for salt and pH, though slightly more stringent conditions may be desired. Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly (e.g. beads placed in scintillant), or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-  
25 PAGE, and the level of *RAFTK*-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques.

Other techniques for immobilizing proteins on matrices are also available for use in the subject assay. For instance, either *RAFTK* or its cognate binding protein can be immobilized utilizing conjugation of biotin and streptavidin. For instance,  
30 biotinylated *RAFTK* molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with *RAFTK* but which do not interfere with binding of upstream or downstream elements can be derivatized  
35 to the wells of the plate, and *RAFTK* trapped in the wells by antibody conjugation. As above, preparations of a *RAFTK*-binding protein and a test compound are incubated in the *RAFTK*-presenting wells of the plate, and the amount of complex trapped in the

well can be quantitated. Exemplary methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the *RAFTK* binding element, or which are reactive with *RAFTK* protein and compete with the binding element; as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the binding element, either intrinsic or extrinsic activity. In the instance of the latter, the enzyme can be chemically conjugated or provided as a fusion protein with the *RAFTK*-BP. To illustrate, the *RAFTK*-BP can be chemically cross-linked or genetically fused with horseradish peroxidase, and the amount of polypeptide trapped in the complex can be assessed with a chromogenic substrate of the enzyme, e.g. 3,3'-diamino-benzadine tetrahydrochloride or 4-chloro-1-naphthol. Likewise, a fusion protein comprising the polypeptide and glutathione-S-transferase can be provided, and complex formation quantitated by detecting the GST activity using 1-chloro-2,4-dinitrobenzene (Habig et al (1974) *J Biol Chem* 249, 7130).

For processes which rely on immunodetection for quantitating one of the proteins trapped in the complex, antibodies against the protein, such as anti-*RAFTK* antibodies, can be used. Alternatively, the protein to be detected in the complex can be "epitope tagged" in the form of a fusion protein which includes, in addition to the *RAFTK* sequence, a second polypeptide for which antibodies are readily available (e.g. from commercial sources). For instance, the GST fusion proteins described above can also be used for quantification of binding using antibodies against the GST moiety. Other useful epitope tags include myc-epitopes (e.g., see Ellison et al. (1991) *J Biol Chem* 266, 21150-21157) which includes a 10-residue sequence from c-myc, as well as the pFLAG system (International Biotechnologies, Inc.) or the pEZZ-protein A system (Pharmacia, NJ).

#### *Cell based assays*

In addition to cell-free assays, such as described above, the readily available source of mammalian *RAFTK* proteins provided by the present invention also facilitates the generation of cell-based assays for identifying small molecule agonists/antagonists and the like. For example, cells which are sensitive to ligands which transduce signals via a pathway involving *RAFTK* can be caused to overexpress a recombinant *RAFTK* protein in the presence and absence of a test agent of interest, with the assay scoring for modulation of *RAFTK* responses by the target cell mediated by the test agent. As with the cell-free assays, agents which produce a statistically significant change in *RAFTK*-dependent responses (either inhibition or potentiation) can be identified. In an illustrative embodiment, the expression or

activity of a *RAFTK* is modulated in cells and the effects of compounds of interest on the readout of interest (such as apoptosis, proliferation or differentiation) are measured. For example, the expression of genes which are up- or down-regulated in response to a *RAFTK*-dependent signal cascade can be assayed. In preferred  
5       embodiments, the regulatory regions of such genes, e.g., the 5' flanking promoter and enhancer regions, are operably linked to a detectable marker (such as luciferase) which encodes a gene product that can be readily detected. Phosphorylation of *RAFTK* or *RAFTK* binding proteins can also be measured, for example by immunoblotting as described in the appended examples.

10       Monitoring the influence of compounds on cells may be applied not only in basic drug screening, but also in clinical trials. In such clinical trials, the expression of a panel of genes may be used as a "read out" of a particular drug's therapeutic effect.

15       In yet another aspect of the invention, the subject *RAFTK* polypeptides can be used to generate a "two hybrid" assay (see, for example, U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72, 223-232; Madura et al. (1993) *J Biol Chem* 268, 12046-12054; Bartel et al. (1993) *Biotechniques* 14, 920-924; Iwabuchi et al. (1993) *Oncogene* 8, 1693-1696; and Brent WO94/10300), for isolating coding sequences for other cellular proteins which bind to or interact with *RAFTK* ("*RAFTK*-binding  
20       proteins" or "*RAFTK*-bp". Such *RAFTK*-binding proteins would likely also be involved in the propagation of signals by the *RAFTK* proteins as, for example, the upstream or downstream elements of the *RAFTK* pathway.

25       Briefly, the two hybrid assay relies on reconstituting in vivo a functional transcriptional activator protein from two separate fusion proteins. In particular, the method makes use of chimeric genes which express hybrid proteins. To illustrate, a first hybrid gene comprises the coding sequence for a DNA-binding domain of a transcriptional activator fused in frame to the coding sequence for a *RAFTK* polypeptide. The second hybrid protein encodes a transcriptional activation domain fused in frame to a sample gene from a cDNA library. If the bait and sample hybrid  
30       proteins are able to interact, e.g., form a *RAFTK*-dependent complex, they bring into close proximity the two domains of the transcriptional activator. This proximity is sufficient to cause transcription of a reporter gene which is operably linked to a transcriptional regulatory site responsive to the transcriptional activator, and expression of the reporter gene can be detected and used to score for the interaction of  
35       the *RAFTK* and sample proteins.

*Transgenic animals*

These systems may be used in a variety of applications. For example, the cell- and animal-based model systems may be used to further characterize *RAFTK* genes and proteins. In addition, such assays may be utilized as part of screening strategies designed to identify compounds which are capable of ameliorating disease symptoms. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions which may be effective in treating disease.

One aspect of the present invention concerns transgenic animals comprising of cells which contain a transgene of the present invention and which preferably (though optionally) express an exogenous *RAFTK* protein in one or more cells in the animal. A *RAFTK* transgene can encode the wild-type form of the protein, or can encode homologs thereof, including both agonists and antagonists, as well as antisense constructs. In preferred embodiments, the expression of the transgene is restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, cis-acting sequences that control expression in the desired pattern. In the present invention, such mosaic expression of a *RAFTK* protein can be essential for many forms of lineage analysis and can additionally provide a means to assess the effects of, for example, lack of *RAFTK* expression which might grossly alter development in small patches of tissue within an otherwise normal embryo. Toward this end, tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences.

Genetic techniques which allow for the expression of transgenes can be regulated via site-specific genetic manipulation in vivo are known to those skilled in the art. For instance, genetic systems are available which allow for the regulated expression of a recombinase that catalyzes the genetic recombination a target sequence. As used herein, the phrase "target sequence" refers to a nucleotide sequence that is genetically recombined by a recombinase. The target sequence is flanked by recombinase recognition sequences and is generally either excised or inverted in cells expressing recombinase activity. Recombinase catalyzed recombination events can be designed such that recombination of the target sequence results in either the activation or repression of expression of one of the subject *RAFTK* proteins. For example, excision of a target sequence which interferes with the expression of a recombinant *RAFTK* gene, such as one which encodes an antagonistic homolog or an antisense transcript, can be designed to activate expression of that



gene. This interference with expression of the protein can result from a variety of mechanisms, such as spatial separation of the *RAFTK* gene from the promoter element or an internal stop codon. Moreover, the transgene can be made wherein the coding sequence of the gene is flanked by recombinase recognition sequences and is initially  
5 transfected into cells in a 3' to 5' orientation with respect to the promoter element. In such an instance, inversion of the target sequence will reorient the subject gene by placing the 5' end of the coding sequence in an orientation with respect to the promoter element which allow for promoter driven transcriptional activation.

The transgenic animals of the present invention all include within a plurality  
10 of their cells a transgene of the present invention, which transgene alters the phenotype of the "host cell" with respect to regulation of cell growth, death and/or differentiation. Since it is possible to produce transgenic organisms of the invention utilizing one or more of the transgene constructs described herein, a general description will be given of the production of transgenic organisms by referring  
15 generally to exogenous genetic material. This general description can be adapted by those skilled in the art in order to incorporate specific transgene sequences into organisms utilizing the methods and materials described below.

In an illustrative embodiment, either the cre/loxP recombinase system of bacteriophage P1 (Lakso et al. (1992) *PNAS* 89, 6232-6236; Orban et al. (1992) *PNAS*  
20 89, 6861-6865) or the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) *Science* 251, 1351-1355; PCT publication WO 92/15694) can be used to generate in vivo site-specific genetic recombination systems. Cre recombinase catalyzes the site-specific recombination of an intervening target sequence located between loxP sequences. loxP sequences are 34 base pair nucleotide  
25 repeat sequences to which the Cre recombinase binds and are required for Cre recombinase mediated genetic recombination. The orientation of loxP sequences determines whether the intervening target sequence is excised or inverted when Cre recombinase is present (Abremski et al. (1984) *J. Biol. Chem.* 259, 1509-1514); catalyzing the excision of the target sequence when the loxP sequences are oriented as  
30 direct repeats and catalyzes inversion of the target sequence when loxP sequences are oriented as inverted repeats.

Accordingly, genetic recombination of the target sequence is dependent on expression of the Cre recombinase. Expression of the recombinase can be regulated by promoter elements which are subject to regulatory control, e.g., tissue-specific,  
35 developmental stage-specific, inducible or repressible by externally added agents. This regulated control will result in genetic recombination of the target sequence only in cells where recombinase expression is mediated by the promoter element. Thus,

the activation expression of a recombinant *RAFTK* protein can be regulated via control of recombinase expression.

Use of the cre/loxP recombinase system to regulate expression of a recombinant *RAFTK* protein requires the construction of a transgenic animal containing transgenes encoding both the Cre recombinase and the subject protein. Animals containing both the Cre recombinase and a recombinant *RAFTK* gene can be provided through the construction of "double" transgenic animals. A convenient method for providing such animals is to mate two transgenic animals each containing a transgene, e.g., a *RAFTK* gene and recombinase gene.

One advantage derived from initially constructing transgenic animals containing a *RAFTK* transgene in a recombinase-mediated expressible format derives from the likelihood that the subject protein, whether agonistic or antagonistic, can be deleterious upon expression in the transgenic animal. In such an instance, a founder population, in which the subject transgene is silent in all tissues, can be propagated and maintained. Individuals of this founder population can be crossed with animals expressing the recombinase in, for example, one or more tissues and/or a desired temporal pattern. Thus, the creation of a founder population in which, for example, an antagonistic *RAFTK* transgene is silent will allow the study of progeny from that founder in which disruption of *RAFTK* mediated induction in a particular tissue or at certain developmental stages would result in, for example, a lethal phenotype.

Similar conditional transgenes can be provided using prokaryotic promoter sequences which require prokaryotic proteins to be simultaneously expressed in order to facilitate expression of the *RAFTK* transgene. Exemplary promoters and the corresponding trans-activating prokaryotic proteins are given in U.S. Patent No. 4,833,080.

Moreover, expression of the conditional transgenes can be induced by gene therapy-like methods wherein a gene encoding the trans-activating protein, e.g. a recombinase or a prokaryotic protein, is delivered to the tissue and caused to be expressed, such as in a cell-type specific manner. By this method, a *RAFTK* transgene could remain silent into adulthood until "turned on" by the introduction of the trans-activator.

In an exemplary embodiment, the "transgenic non-human animals" of the invention are produced by introducing transgenes into the germline of the non-human animal. Embryonal target cells at various developmental stages can be used to introduce transgenes. Different methods are used depending on the stage of development of the embryonal target cell. The specific line(s) of any animal used to practice this invention are selected for general good health, good embryo yields, good

pronuclear visibility in the embryo, and good reproductive fitness. In addition, the haplotype is a significant factor. For example, when transgenic mice are to be produced, strains such as C57BL/6 or FVB lines are often used (Jackson Laboratory, Bar Harbor, ME). Preferred strains are those with H-2b, H-2d or H-2q haplotypes  
5 such as C57BL/6 or DBA/1. The line(s) used to practice this invention may themselves be transgenics, and/or may be knockouts (i.e., obtained from animals which have one or more genes partially or completely suppressed) .

In one embodiment, the transgene construct is introduced into a single stage embryo. The zygote is the best target for micro-injection. In the mouse, the male  
10 pronucleus reaches the size of approximately 20 micrometers in diameter which allows reproducible injection of 1-2pl of DNA solution. The use of zygotes as a target for gene transfer has a major advantage in that in most cases the injected DNA will be incorporated into the host gene before the first cleavage (Brinster et al. (1985) *PNAS* 82, 4438-4442). As a consequence, all cells of the transgenic animal will carry the  
15 incorporated transgene. This will in general also be reflected in the efficient transmission of the transgene to offspring of the founder since 50% of the germ cells will harbor the transgene.

Normally, fertilized embryos are incubated in suitable media until the pronuclei appear. At about this time, the nucleotide sequence comprising the  
20 transgene is introduced into the female or male pronucleus as described below. In some species such as mice, the male pronucleus is preferred. It is most preferred that the exogenous genetic material be added to the male DNA complement of the zygote prior to its being processed by the ovum nucleus or the zygote female pronucleus. It is thought that the ovum nucleus or female pronucleus release molecules which affect  
25 the male DNA complement, perhaps by replacing the protamines of the male DNA with histones, thereby facilitating the combination of the female and male DNA complements to form the diploid zygote.

Thus, it is preferred that the exogenous genetic material be added to the male complement of DNA or any other complement of DNA prior to its being affected by  
30 the female pronucleus. For example, the exogenous genetic material is added to the early male pronucleus, as soon as possible after the formation of the male pronucleus, which is when the male and female pronuclei are well separated and both are located close to the cell membrane. Alternatively, the exogenous genetic material could be added to the nucleus of the sperm after it has been induced to undergo  
35 decondensation. Sperm containing the exogenous genetic material can then be added to the ovum or the decondensed sperm could be added to the ovum with the transgene constructs being added as soon as possible thereafter.

Introduction of the transgene nucleotide sequence into the embryo may be accomplished by any means known in the art such as, for example, microinjection, electroporation, or lipofection. Following introduction of the transgene nucleotide sequence into the embryo, the embryo may be incubated in vitro for varying amounts  
5 of time, or reimplanted into the surrogate host, or both. In vitro incubation to maturity is within the scope of this invention. One common method in to incubate the embryos in vitro for about 1-7 days, depending on the species, and then reimplant them into the surrogate host.

For the purposes of this invention a zygote is essentially the formation of a  
10 diploid cell which is capable of developing into a complete organism. Generally, the zygote will be comprised of an egg containing a nucleus formed, either naturally or artificially, by the fusion of two haploid nuclei from a gamete or gametes. Thus, the gamete nuclei must be ones which are naturally compatible, i.e., ones which result in a  
15 viable zygote capable of undergoing differentiation and developing into a functioning organism. Generally, a euploid zygote is preferred. If an aneuploid zygote is obtained, then the number of chromosomes should not vary by more than one with respect to the euploid number of the organism from which either gamete originated.

In addition to similar biological considerations, physical ones also govern the amount (e.g., volume) of exogenous genetic material which can be added to the  
20 nucleus of the zygote or to the genetic material which forms a part of the zygote nucleus. If no genetic material is removed, then the amount of exogenous genetic material which can be added is limited by the amount which will be absorbed without being physically disruptive. Generally, the volume of exogenous genetic material  
25 inserted will not exceed about 10 picoliters. The physical effects of addition must not be so great as to physically destroy the viability of the zygote. The biological limit of the number and variety of DNA sequences will vary depending upon the particular zygote and functions of the exogenous genetic material and will be readily apparent to one skilled in the art, because the genetic material, including the exogenous genetic  
30 material, of the resulting zygote must be biologically capable of initiating and maintaining the differentiation and development of the zygote into a functional organism.

The number of copies of the transgene constructs which are added to the zygote is dependent upon the total amount of exogenous genetic material added and will be the amount which enables the genetic transformation to occur. Theoretically  
35 only one copy is required; however, generally, numerous copies are utilized, for example, 1,000-20,000 copies of the transgene construct, in order to insure that one copy is functional. As regards the present invention, there will often be an advantage

to having more than one functioning copy of each of the inserted exogenous DNA sequences to enhance the phenotypic expression of the exogenous DNA sequences.

Any technique which allows for the addition of the exogenous genetic material into nucleic genetic material can be utilized so long as it is not destructive to the cell, nuclear membrane or other existing cellular or genetic structures. The exogenous genetic material is preferentially inserted into the nucleic genetic material by microinjection. Microinjection of cells and cellular structures is known and is used in the art.

Reimplantation is accomplished using standard methods. Usually, the surrogate host is anesthetized, and the embryos are inserted into the oviduct. The number of embryos implanted into a particular host will vary by species, but will usually be comparable to the number of off spring the species naturally produces.

Transgenic offspring of the surrogate host may be screened for the presence and/or expression of the transgene by any suitable method. Screening is often accomplished by Southern blot or Northern blot analysis, using a probe that is complementary to at least a portion of the transgene. Western blot analysis using an antibody against the protein encoded by the transgene may be employed as an alternative or additional method for screening for the presence of the transgene product. Typically, DNA is prepared from tail tissue and analyzed by Southern analysis or PCR for the transgene. Alternatively, the tissues or cells believed to express the transgene at the highest levels are tested for the presence and expression of the transgene using Southern analysis or PCR, although any tissues or cell types may be used for this analysis.

Alternative or additional methods for evaluating the presence of the transgene include, without limitation, suitable biochemical assays such as enzyme and/or immunological assays, histological stains for particular marker or enzyme activities, flow cytometric analysis, and the like. Analysis of the blood may also be useful to detect the presence of the transgene product in the blood, as well as to evaluate the effect of the transgene on the levels of various types of blood cells and other blood constituents.

Progeny of the transgenic animals may be obtained by mating the transgenic animal with a suitable partner, or by *in vitro* fertilization of eggs and/or sperm obtained from the transgenic animal. Where mating with a partner is to be performed, the partner may or may not be transgenic and/or a knockout; where it is transgenic, it may contain the same or a different transgene, or both. Alternatively, the partner may be a parental line. Where *in vitro* fertilization is used, the fertilized embryo may be

implanted into a surrogate host or incubated in vitro, or both. Using either method, the progeny may be evaluated for the presence of the transgene using methods described above, or other appropriate methods.

The transgenic animals produced in accordance with the present invention will include exogenous genetic material. As set out above, the exogenous genetic material will, in certain embodiments, be a DNA sequence which results in the production of a *RAFTK* protein (either agonistic or antagonistic), and antisense transcript, or a *RAFTK* mutant. Further, in such embodiments the sequence will be attached to a transcriptional control element, e.g., a promoter, which preferably allows the expression of the transgene product in a specific type of cell.

Retroviral infection can also be used to introduce transgene into a non-human animal. The developing non-human embryo can be cultured in vitro to the blastocyst stage. During this time, the blastomeres can be targets for retroviral infection (Jaenich, R. (1976) *PNAS* 73, 1260-1264). Efficient infection of the blastomeres is obtained by enzymatic treatment to remove the zona pellucida (*Manipulating the Mouse Embryo*, Hogan eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1986). The viral vector system used to introduce the transgene is typically a replication-defective retrovirus carrying the transgene (Jahner et al. (1985) *PNAS* 82, 6927-6931; Van der Putten et al. (1985) *PNAS* 82, 6148-6152). Transfection is easily and efficiently obtained by culturing the blastomeres on a monolayer of virus-producing cells (Van der Putten, supra; Stewart et al. (1987) *EMBO J.* 6, 383-388). Alternatively, infection can be performed at a later stage. Virus or virus-producing cells can be injected into the blastocoele (Jahner et al. (1982) *Nature* 298, 623-628). Most of the founders will be mosaic for the transgene since incorporation occurs only in a subset of the cells which formed the transgenic non-human animal. Further, the founder may contain various retroviral insertions of the transgene at different positions in the genome which generally will segregate in the offspring. In addition, it is also possible to introduce transgenes into the germ line by intrauterine retroviral infection of the midgestation embryo (Jahner et al. (1982) supra).

A third type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells are obtained from pre-implantation embryos cultured in vitro and fused with embryos (Evans et al. (1981) *Nature* 292, 154-156; Bradley et al. (1984) *Nature* 309, 255-258; Gossler et al. (1986) *PNAS* 83, 9065-9069; and Robertson et al. (1986) *Nature* 322, 445-448). Transgenes can be efficiently introduced into the ES cells by DNA transfection or by retrovirus-mediated transduction. Such transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The ES cells

thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal. For review see Jaenisch, R. (1988) *Science* 240, 1468-1474.

In one embodiment, gene targeting, which is a method of using homologous recombination to modify an animal's genome, can be used to introduce changes into cultured embryonic stem cells. By targeting a *RAFTK* gene of interest in ES cells, these changes can be introduced into the germlines of animals to generate chimeras. The gene targeting procedure is accomplished by introducing into tissue culture cells a DNA targeting construct that includes a segment homologous to a target *RAFTK* locus, and which also includes an intended sequence modification to the *RAFTK* genomic sequence (e.g., insertion, deletion, point mutation). The treated cells are then screened for accurate targeting to identify and isolate those which have been properly targeted.

Gene targeting in embryonic stem cells is in fact a scheme contemplated by the present invention as a means for disrupting a *RAFTK* gene function through the use of a targeting transgene construct designed to undergo homologous recombination with one or more *RAFTK* genomic sequences. The targeting construct can be arranged so that, upon recombination with an element of a *RAFTK* gene, a positive selection marker is inserted into (or replaces) coding sequences of the targeted *RAFTK* gene. The inserted sequence functionally disrupts the *RAFTK* gene, while also providing a positive selection trait. Exemplary *RAFTK* targeting constructs are described in more detail below.

Generally, the embryonic stem cells (ES cells) used to produce the knockout animals will be of the same species as the knockout animal to be generated. Thus for example, mouse embryonic stem cells will usually be used for generation of knockout mice.

Embryonic stem cells are generated and maintained using methods well known to the skilled artisan such as those described by Doetschman et al. (1985) *J. Embryol. Exp. Morphol.* 87, 27-45). Any line of ES cells can be used, however, the line chosen is typically selected for the ability of the cells to integrate into and become part of the germ line of a developing embryo so as to create germ line transmission of the knockout construct. Thus, any ES cell line that is believed to have this capability is suitable for use herein. One mouse strain that is typically used for production of ES cells, is the 129J strain. Another ES cell line is murine cell line D3 (American Type Culture Collection, catalog no. CKL 1934) Still another preferred ES cell line is the WW6 cell line (Ioffe et al. (1995) *PNAS* 92, 7357-7361). The cells are cultured and prepared for knockout construct insertion using methods well known to the skilled artisan, such as those set forth by Robertson in: *Teratocarcinomas and Embryonic*

*Stem Cells: A Practical Approach*, E.J. Robertson, ed. IRL Press, Washington, D.C. [1987]); by Bradley et al. (1986) *Current Topics in Devel. Biol.* 20, 357-371); and by Hogan et al. (*Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY [1986]) .

5           Insertion of the knockout construct into the ES cells can be accomplished using a variety of methods well known in the art including for example, electroporation, microinjection, and calcium phosphate treatment. A preferred method of insertion is electroporation .

10           Each knockout construct to be inserted into the cell must first be in the linear form. Therefore, if the knockout construct has been inserted into a vector (described infra), linearization is accomplished by digesting the DNA with a suitable restriction endonuclease selected to cut only within the vector sequence and not within the knockout construct sequence.

15           For insertion, the knockout construct is added to the ES cells under appropriate conditions for the insertion method chosen, as is known to the skilled artisan. Where more than one construct is to be introduced into the ES cell, each knockout construct can be introduced simultaneously or one at a time.

20           If the ES cells are to be electroporated, the ES cells and knockout construct DNA are exposed to an electric pulse using an electroporation machine and following the manufacturer's guidelines for use. After electroporation, the ES cells are typically allowed to recover under suitable incubation conditions. The cells are then screened for the presence of the knockout construct.

25           Screening can be accomplished using a variety of methods. Where the marker gene is an antibiotic resistance gene, for example, the ES cells may be cultured in the presence of an otherwise lethal concentration of antibiotic. Those ES cells that survive have presumably integrated the knockout construct. If the marker gene is other than an antibiotic resistance gene, a Southern blot of the ES cell genomic DNA can be probed with a sequence of DNA designed to hybridize only to the marker sequence. Alternatively, PCR can be used. Finally, if the marker gene is a gene that encodes an  
30           enzyme whose activity can be detected (e.g., b-galactosidase), the enzyme substrate can be added to the cells under suitable conditions, and the enzymatic activity can be analyzed. One skilled in the art will be familiar with other useful markers and the means for detecting their presence in a given cell. All such markers are contemplated as being included within the scope of the teaching of this invention.

35           The knockout construct may integrate into several locations in the ES cell genome, and may integrate into a different location in each ES cell's genome due to the occurrence of random insertion events. The desired location of insertion is in a



complementary position to the DNA sequence to be knocked out, e.g., the *RAFTK* coding sequence, transcriptional regulatory sequence, etc. Typically, less than about 1-5 % of the ES cells that take up the knockout construct will actually integrate the knockout construct in the desired location. To identify those ES cells with proper  
5 integration of the knockout construct, total DNA can be extracted from the ES cells using standard methods. The DNA can then be probed on a Southern blot with a probe or probes designed to hybridize in a specific pattern to genomic DNA digested with particular restriction enzyme(s). Alternatively, or additionally, the genomic DNA can be amplified by PCR with probes specifically designed to amplify DNA  
10 fragments of a particular size and sequence (i.e., only those cells containing the knockout construct in the proper position will generate DNA fragments of the proper size).

After suitable ES cells containing the knockout construct in the proper location have been identified, the cells can be inserted into an embryo. Insertion may be  
15 accomplished in a variety of ways known to the skilled artisan, however a preferred method is by microinjection. For microinjection, about 10-30 cells are collected into a micropipet and injected into embryos that are at the proper stage of development to permit integration of the foreign ES cell containing the knockout construct into the developing embryo. For instance, as the appended Examples describe, the  
20 transformed ES cells can be microinjected into blastocytes.

The suitable stage of development for the embryo used for insertion of ES cells is very species dependent, however for mice it is about 3.5 days. The embryos are obtained by perfusing the uterus of pregnant females. Suitable methods for accomplishing this are known to the skilled artisan, and are set forth by, e.g., Bradley  
25 et al. (*supra*).

While any embryo of the right stage of development is suitable for use, preferred embryos are male. In mice, the preferred embryos also have genes coding for a coat color that is different from the coat color encoded by the ES cell genes. In this way, the offspring can be screened easily for the presence of the knockout  
30 construct by looking for mosaic coat color (indicating that the ES cell was incorporated into the developing embryo). Thus, for example, if the ES cell line carries the genes for white fur, the embryo selected will carry genes for black or brown fur.

After the ES cell has been introduced into the embryo, the embryo may be  
35 implanted into the uterus of a pseudopregnant foster mother for gestation. While any foster mother may be used, the foster mother is typically selected for her ability to breed and reproduce well, and for her ability to care for the young. Such foster

mothers are typically prepared by mating with vasectomized males of the same species. The stage of the pseudopregnant foster mother is important for successful implantation, and it is species dependent. For mice, this stage is about 2-3 days pseudopregnant.

5        Offspring that are born to the foster mother may be screened initially for mosaic coat color where the coat color selection strategy (as described above, and in the appended examples) has been employed. In addition, or as an alternative, DNA from tail tissue of the offspring may be screened for the presence of the knockout construct using Southern blots and/or PCR as described above. Offspring that appear  
10       to be mosaics may then be crossed to each other, if they are believed to carry the knockout construct in their germ line, in order to generate homozygous knockout animals. Homozygotes may be identified by Southern blotting of equivalent amounts of genomic DNA from mice that are the product of this cross, as well as mice that are known heterozygotes and wild type mice.

15       Other means of identifying and characterizing the knockout offspring are available. For example, Northern blots can be used to probe the mRNA for the presence or absence of transcripts encoding either the gene knocked out, the marker gene, or both. In addition, Western blots can be used to assess the level of expression of the *RAFTK* gene knocked out in various tissues of the offspring by probing the  
20       Western blot with an antibody against the particular *RAFTK* protein, or an antibody against the marker gene product, where this gene is expressed. Finally, in situ analysis (such as fixing the cells and labeling with antibody) and/or FACS (fluorescence activated cell sorting) analysis of various cells from the offspring can be conducted using suitable antibodies to look for the presence or absence of the knockout construct  
25       gene product.

      Yet other methods of making knock-out or disruption transgenic animals are also generally known. See, for example, *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Recombinase dependent knockouts can also be generated, e.g. by homologous recombination to  
30       insert target sequences, such that tissue specific and/or temporal control of inactivation of a *RAFTK*-gene can be controlled by recombinase sequences (described infra).

      Animals containing more than one knockout construct and/or more than one transgene expression construct are prepared in any of several ways. The preferred  
35       manner of preparation is to generate a series of mammals, each containing one of the desired transgenic phenotypes. Such animals are bred together through a series of crosses, backcrosses and selections, to ultimately generate a single animal containing

all desired knockout constructs and/or expression constructs, where the animal is otherwise congenic (genetically identical) to the wild type except for the presence of the knockout construct(s) and/or transgene(s).

The present invention is further illustrated by the following examples which should not be construed as limiting in any way. The contents of all cited references, including literature references, issued patents, published patent applications as cited throughout this application are hereby expressly incorporated by reference. The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning, Volumes I and II* (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al. U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology, Volumes I-IV* (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

**EXAMPLE 1: Identification and Characterization of a Novel Related Adhesion Focal Tyrosine Kinase (RAFTK) from Megakaryocytes and Brain**

30

A cDNA encoding a novel human intracytoplasmic tyrosine kinase, termed *RAFTK* (for a related adhesion focal tyrosine kinase) was isolated. The murine homolog of the human *RAFTK* cDNA was also cloned and characterized.

Comparison of the deduced amino acid sequences of human and murine *RAFTK* cDNAs revealed 95% homology, indicating that *RAFTK* is highly conserved between these species. The *RAFTK* cDNA clone, encoding a polypeptide of 1009 amino acids, has closest homology (48% identity, 65% similarity) to the focal adhesion kinase

35

(pp125<sup>FAK</sup>). Comparison of the deduced amino acid sequences also indicates that *RAFTK* like pp125<sup>FAK</sup> lacks a transmembrane region, myristylation sites and SH2 and SH3 domains. In addition, like pp125<sup>FAK</sup>, *RAFTK* contains a kinase domain flanked by large N-terminal (426 residues) and C-terminal (331 residues) domains, and the C-terminal region contains a predicted proline-rich stretch of residues. In fetal tissues, *RAFTK* expression was abundant in brain and low levels were observed in lung and liver. In adult tissues, it was less restricted indicating that *RAFTK* expression is developmentally upregulated. Expression of *RAFTK* was also observed in human CD34<sup>+</sup> marrow cells, primary bone marrow megakaryocytes, platelets and various areas of brain. The human *RAFTK* gene was assigned to human Chromosome 8 using genomic DNAs from human/rodent somatic cell hybrid lines. The mouse *RAFTK* gene was mapped to Chromosome 14 closely linked to gonadotropin releasing hormone. Using specific antibodies for *RAFTK*, an approximately 123 Kd protein from the human CMK megakaryocyte cell line was immunoprecipitated. Treatment of the CMK megakaryocytic cells with thrombin caused a rapid induction of tyrosine phosphorylation of *RAFTK* protein. The structural features of *RAFTK* suggest that it is a member of the focal adhesion kinase gene family and may participate in signal transduction in human megakaryocytes and brain as well as other cell types.

The predicted amino acid sequence of the *RAFTK* protein shares consensus motifs in the central catalytic domain common to protein tyrosine kinases. The *RAFTK* cDNA, encoding a polypeptide of 1009 amino acids, has the closest homology (48% identity, 65% similarity) to FAK. Analysis of their deduced amino acid sequences also indicates that *RAFTK*, like FAK, lacks a transmembrane region, myristylation sites, and SH2 and SH3 domains. In addition, like FAK, the *RAFTK* C-terminal domain contains a predicted proline-rich stretch of residues. *RAFTK* was reported to be highly expressed in the central nervous system (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10) and involved in Ca<sup>2+</sup>-induced regulation of ion channel and MAP kinase functions in PC-12 cells (Lev, S. et al. (1995) *Nature* 376, 737).

*RAFTK* expression is abundant in primary bone marrow megakaryocytes and their progeny, blood platelets (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10). To address the role of *RAFTK* in signal transduction pathways in megakaryocytes, experiments were performed using the model CMK megakaryocytic cell line (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485). The c-kit receptor and its cognate ligand SCF were investigated since they play a critical role in the adhesion, migration, motility, proliferation and maturation of a number of hematopoietic cells, including megakaryocytes and platelets (See, e.g., Dastych, J. and Metcalfe, D. D. (1994) *J.*

*Immunol.* 152, 213-219; Kinashi, T. et al. (1995) *Blood* 86, 2086-2090; Scott, G. et al. (1994) *Pigment Cell Res.* 7, 44-51). Since SCF effects appear to be modulated by PKC (Blume-Jensen, P. et al. (1995) *Journal of Biological Chemistry* 270, 14192-14200), the ability of PKC to mediate the effects of SCF and  $Ca^{2+}$  on *RAFTK* phosphorylation was investigated. In this study, SCF and PMA induced the tyrosine phosphorylation of *RAFTK* through PKC. In addition, *RAFTK* was associated with the cytoskeletal protein paxillin in megakaryocytes, and this association appeared critical for *RAFTK* phosphorylation.

10           The following materials and methods were used to clone and characterize *RAFTK*

### Materials

15           Chemical reagents were purchased from Sigma (St. Louis, MO). Restriction endonucleases, modifying enzymes, and terminal deoxynucleotidyl transferase were purchased from Pharmacia Biotech, Inc. (Piscataway, NJ) and New England Biolabs (Beverly, MA). The primers for polymerase chain reaction (PCR), RNA-PCR and sequencing were synthesized by an automated DNA synthesizer (Applied Biosystems, model 394). The PCR and RNA-PCR reagents were obtained from Perkin-Elmer  
20           Cetus (Norwalk, CT) and random primed labeling kits were obtained from Stratagene (La Jolla, CA). Manual and automated sequencing kits were obtained from USB (Cleveland, OH) and Pharmacia Biotech, Inc., respectively. Automated sequencing was performed using Pharmacia's Automated Laser Fluorescent Sequencer (ALF). Monoclonal antibody 2A7 against pp125<sup>FAK</sup> protein was kindly obtained from Dr. J.  
25           Thomas Parsons (Charlottesville, VA). Monoclonal antibody PY-20 directed against Tyr(P) was obtained from ICN (Costa Mesa, CA).

### Cells

30           Human marrow megakaryocytes were isolated by a method employing immunomagnetic beads using anti-human glycoprotein GpIIb monoclonal antibody, as previously described (Tanaka, H. et al. (1989) *Br. J. Haematol.* 73, 18-22; Avraham, H. et al. (1992) *Blood* 80, 1679-1684). CD34 bearing marrow progenitor cells were purified from heparinized bone marrow aspirates using immunomagnetic beads coated with an anti-CD34 monoclonal antibody as previously described  
35           (Avraham, H. et al. (1992) *Blood* 80, 1679-1684). The CMK cell line, provided by Dr. T. Sato and derived from a child with megakaryoblastic leukemia, has properties of cells of the megakaryocytic lineage (Sato, T. et al. (1987) *Exp. Hematol.* (N.Y.) 15,

495-502). The CMK cell line was cultured in RPMI 1640 medium with 10% fetal calf serum. Additional permanent human megakaryocytic cell lines were studied. DAMI cells were obtained from Dr. S. Greenberg, (Brigham and Women's Hospital, Boston, MA), Mo7e and erythroid-megakaryocytic HEL cells were obtained from Dr. L. Zon, (Children's Hospital, Boston, MA). Each cell line was cultured as previously described (Avraham, H. et al. (1992) *Blood* 80, 1679-1684; Avraham, H. et al. (1992) *Blood* 79, 365-371; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-75). Other permanent human cell lines such as Ramos (human B-cells) were obtained from the American Type Tissue Culture Collection and maintained in liquid culture according to the specifications in the catalog.

Human platelets were isolated by gel filtration from freshly drawn blood anticoagulated with 0.15 vol NIH formula A acid-citrate-dextrose solution supplemented with 1  $\mu$ M prostaglandin E<sub>1</sub> (PGE<sub>1</sub>) as previously described (Lipfert, L. et al. (1992) *J. Cell Biol.* 119, 905-912).

#### DNA amplification and cloning

Total RNA derived from CMK cells was prepared by a standard protocol of lysis in guanidinium isothiocyanate followed by cesium chloride gradient centrifugation (Maniatis, T. et al. (1992) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). Protein-tyrosine kinase sequences were amplified with degenerate oligonucleotide primers as previously described (Wilks, A.F. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1603-1607).

Briefly, total RNA (10  $\mu$ g) was used as a template for synthesis of complementary DNA (cDNA). The PTK3 oligonucleotide "SDVWSF/YG" (SEQ ID NO:5) 5'- (C/G)(T/A)(A/G)TC(A/C/G/T)ACCCA(A/C/G/T)(C/G)(T/A)(A/G)(T/A)A(A/C/G/T)CC - 3' (SEQ ID NO:6) was designed in our lab and was used as a primer. PCR was performed on one quarter of the cDNA synthesis reaction mixture (original volume-20  $\mu$ l), using PTK1 "DLAARN" (SEQ ID NO:7) 5'- CGACGA(T/C)CT(A/C/G/T)GC(A/C/G/T) (A/G)C(A/C/G/T)AA - 3'(SEQ ID NO:8) and PTK2 "WMAPE" (SEQ ID NO:9) 5' - GTACC(T/C)TC(G/C/A)GG(A/C/G/T)GCCATCCA - 3' (SEQ ID NO:10) oligonucleotides (50 pmol each) (Wilks, A.F. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1603-1607). The mixture was then subjected to PCR amplification using the Perkin-Elmer Cetus thermal cycler set for 30 cycles as follows: denature 95°C, 2 min; primer anneal 37°C, 1.5 min; primer extension 72°C, 2.30 min; 1 minute ramp times were used between these temperatures. PCR products of the amplified tyrosine kinase domains were purified from the agarose gel, digested with EcoRI and BamHI, ligated

into pUC19, and transformed into Escherichia coli DH5 $\alpha$ . Sequencing was carried out by the dideoxy chain termination method using version 2.0 sequenase kit (USB, Cleveland, OH). Sequences were compared with known sequences in GenBank and EMBL data bases using the Autosearch computer program. A novel clone was identified. This 160-base pair (bp) PCR product, designated JJ3, was radiolabeled using the "Prime It II" random priming protocol (Stratagene) and used as a probe to screen human cDNA libraries.

#### Isolation and characterization of cDNA clones

The human brain (hippocampus) cDNA library in  $\lambda$ -ZapII vector (randomized and oligo dT, cat #936205, Stratagene, CA) was screened ( $\sim 5 \times 10^5$  recombinants/screening) initially with the 160 bp PCR fragment (termed JJ3), and labeled with [ $\gamma^{32}\text{P}$ ] dCTP using random primed cDNA labeling. Hybridization to nylon filters (MSI) was performed in 50% formamide, 6 x SSC, 10 mM sodium phosphate, 5 x Denhardt's solution, 0.1% sodium dodecyl sulfate (SDS), and 1 mg/ml Herring sperm DNA (Boehringer Mannheim, Germany) at 43°C overnight. The filters were washed at room temperature in 2 x SSC, 1% SDS, and then in 0.2 x SSC, 0.1% SDS at 63°C three times for 30 min., UV crosslinked (Stratagene Stratalinker), and exposed to Kodak X-OMAT AR film (Eastman Kodak). Twelve clones were isolated and processed. Plasmid DNA was prepared using Exassist Helper Phage and the SolR System according to the manufacturer's instructions (Stratagene). Of these twelve clones, two were sequenced on both strands. A human CMK-PMA cDNA library oligo dT (Avraham, H. et al. (1992) *Blood* 79, 365-371) ( $\sim 3 \times 10^5$  recombinants/screening) in  $\lambda$ -gt10 vector was screened with the  $^{32}\text{P}$ -labeled JJ3 fragment. Four clones were isolated and the recombinant DNAs of 2 positive phages were digested with EcoRI, and the cDNA insert was subcloned into pBSK (Stratagene) and thereafter sequenced.

A 340 bp probe was prepared from the 5'- end of one of the CMK cDNA clones (termed 2-1) and used to screen the human brain (hippocampus) cDNA library. Twelve clones were isolated and two clones were sequenced on both strands. In addition, a 248 bp probe was prepared from the 5'- end of one of the clones (termed 4C) and the human hippocampus cDNA library was rescreened. Twelve clones were identified and isolated and of these, 1 clone (termed 3B) was sequenced on both strands.

The mouse brain cDNA library (cat # ML1042b, Clontech, Palo Alto, CA) in  $\lambda$ -gt11 vector was screened ( $\sim 5 \times 10^5$  recombinants/screening) using 381 bp 5'- Kpn I fragment or 764 bp ApaI -3'- fragment of human *RAFTK* cDNA as a probe and the

filters were hybridized and washed under high stringency conditions. Six clones were isolated. The DNA was isolated as previously described (Maniatis, T. et al. (1992) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and subcloned into pBSK and thereafter sequenced.

- 5 Nucleotide sequences were determined by the Automated Laser Fluorescent (ALF) DNA sequencer (Pharmacia Biotech, Inc.) using Autoread (Pharmacia) and by manual sequencing using sequenase kit (USB).

### Chromosomal localization of the human *RAFTK* gene

- 10 Genomic DNAs from the NIGMS Hybrid Mapping Panels #1 and #2 were obtained from the NIGMS Genetic Mutant Cell Repository (Coriell Cell Institute for Medical Research, Camden, NJ). In addition, both mapping panels included DNA samples isolated from human and rodent parental cell lines (mouse and Chinese hamster). Approximately 5 µg of DNA from human, hamster, and mouse genomic  
15 DNAs were digested with BamHI, HindIII and PstI to find a suitable restriction fragment length polymorphism (RFLP) or unique genomic fragment for use in mapping. Subsequently, genomic DNAs from each panel were cut with BamHI. Southern blots were probed with a human 1.4 kb *RAFTK* cDNA and hybridizations were carried out as previously described (Rowe, L.B. et al. (1994) *Mamm. Genome* 5, 253-274; White, R.A. et al. (1992) *Nature Genet.* 2, 80-83). Hybrids were scored for  
20 the appropriate human-specific restriction endonuclease fragment on the autoradiographs. The results were compared with the chromosome contents of the hybrid cell lines and the concordance between restriction fragments and specific chromosome content was used to establish the localization of human *RAFTK*.

25

### Backcross mapping of the mouse *RAFTK* gene

- Genomic DNAs from C57BL/6J, Mus spretus and a (M. spretus x C57BL/6J) M. spretus BSS type backcross DNA panel were obtained from The Jackson Laboratory (Bar Harbor, Maine) (Rowe, L.B. et al. (1994) *Mamm. Genome* 5, 253-  
30 274). Southern blots and hybridizations were performed as previously described (White, R.A. et al. (1992) *Nature Genet.* 2, 80-83). Approximately 5 µg of genomic DNAs of C57BL/6J and Mus spretus were digested with 29 different restriction enzymes to identify a potential RFLP genetic marker. The Southern blots were probed with a 1.4 kb human *RAFTK* cDNA fragment labeled with <sup>32</sup>P using a  
35 Decaprime II Kit (Ambion, Inc., Austin, TX). Digestion of the backcross DNA panel with BamHI, Southern blotting and hybridizations were carried out as previously described (White, R.A. et al. (1992) *Nature Genet.* 2, 80-83).



**Recombinant inbred (RI) line mapping of the mouse *RAFTK* gene**

*RAFTK* and *Gnrh* co-segregated in BXD RI lines and mapped to chromosome 14. Genomic DNAs isolated from the progenitors of BXD RI lines (C57BL/6J and DBA/2J) were digested with 29 different restriction enzymes to identify a RFLP genetic marker for mapping. Subsequently, genomic DNAs isolated from the BXD RI lines were digested with *SacI*. Conditions for Southern blots and hybridizations were the same as previously described (White, R.A. et al. (1992) *Nature Genet.* 2, 80-83) and the 1.4 kb human *RAFTK* cDNA was used as a probe. Data were compared with strain distribution patterns (SPDs) recorded in GBASE (1993) (Yang-Feng, T.L. et al. (1986) *Somatic Cell. Mol. Genet.* 12, 95-100).

**Northern blot analysis**

Total RNA was prepared by a standard protocol of lysis in guanidinium isothiocyanate followed by cesium chloride gradient centrifugation (Maniatis, T. et al. (1992) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). The human adult and fetal tissue Northern blots, the brain regions and the human tissue II blots were obtained from Clontech (Palo Alto, CA). Hybridization was carried out according to the manufacturer's instructions. Each RNA blot was probed with a 146 bp, 3'- gene-specific *RAFTK* cDNA radiolabeled to a high specific activity ( $10^8$ - $10^9$  cpm/ $\mu$ g) with [ $\gamma$ - $^{32}$ P]dCTP. The level of expression for each mRNA was also determined densitometrically (EC Apparatus Corp. Densitometer; St. Petersburg, FL). The radioactivity associated with each band was also assayed with a Betascope 603 blot analyzer (Betagen, Mountain View, CA). The same blot was assessed for the presence of the actin or GAPDH specific probes; Actin and GAPDH were used as controls to assure uniform loading.

**PCR blots**

cDNA was prepared from platelets ( $10 \times 10^8$ ), CD34<sup>+</sup> marrow cells ( $10^6$  cells) and bone marrow megakaryocytes ( $10^6$  cells) and amplified by PCR using specific *RAFTK* primers as previously described (Bennett, B.D. et al. (1994) *J. Biol. Chem.* 269, 1068-1074). The sequence of the *RAFTK* upstream primer was 5'-CGGGCCGTGCTGGAGCTCAA - 3' (SEQ ID NO:11)(position 2958 - 2977). The nucleotide sequence of the *RAFTK* downstream primer was 5'-GTCCGTGAAGATGACGGCAA - 3' (SEQ ID NO:12) (position 3084 - 3103). The sequence of the FAK upstream primer was 5'-AAAGCTGTCATCGAGATGTCC -3' (SEQ ID NO:13) (position 2292-2312). The nucleotide sequence of the downstream

primer was 5'- TCGGTGGGTGCTGGCTGGTAGG - 3' (SEQ ID NO:14) (position 2417-2438)(Andre, E., and Becker-Andre, M. (1993) *Biochem. Biophys. Res. Commun.* 190, 140-147). The sequence of the actin upstream primer was 5'- ATCTGGCACCACACCTTCTACAATGAGCTGCG - 3' (SEQ ID NO:15). The  
5 nucleotide sequence of the downstream primer was 5'- CGTCATACTCCTGCTTGCTGATCCACATCTGC -3' (SEQ ID NO:16) (Clontech, Palo Alto, CA). The PCR products were electrophoresed on a 1.5% agarose gel, denatured, neutralized, transferred to filters, and vacuum blotted. The probes used were the *RAFTK*, FAK and actin gene-specific probes, which were labeled by random  
10 priming as described above. Prehybridization and hybridization were carried out as previously described (Bennett, B.D. et al. (1994) *J. Biol. Chem.* 269, 1068-1074).

### Protein analysis

Metabolic labeling, immunoprecipitation, and Western blot analysis were  
15 performed in CMK cells as previously described (Laemmli, U.K. (1970) *Nature* 227, 680-685; Yarden, Y et al. (1987) *EMBO J.* 6, 3341-3351; Konopka, J. B., and Witte, O.N. (1985) *Mol. Cell Biol.* 5, 3116-3123; Konopka, J. B. et al. (1984) *J. Virol.* 51, 223-232). For immunoblot analysis, total cell lysates of CMK cells untreated or stimulated with  $\alpha$ -thrombin (1 U/ml or 2 U/ml as indicated)(ChromoLog Corp.,  
20 Havertown, PA) for 5 min were prepared as previously described (Yarden, Y et al. (1987) *EMBO J.* 6, 3341-3351). Relative protein concentrations were determined with a colorimetric assay kit (Bio-Rad Laboratories, Inc., Hercules, CA) with bovine serum albumin as the standard. A portion of lysate containing approximately 0.05 mg of protein was mixed with an equal volume of 2 x SDS sample buffer containing  $\beta$ -  
25 mercaptoethanol, boiled for 5 min., fractionated on 8% polyacrylamide-SDS gels (Laemmli, U.K. (1970) *Nature* 227, 680-685) and transferred to Immobilon polyvinylidene difluoride (Millipore Corp., Bedford, MA) filters. Protein blots were treated with specific *RAFTK* antibodies (R-4250) (see below). Primary binding of the *RAFTK* antibodies (see below) was detected using anti-IgG second antibodies  
30 conjugated to horseradish peroxidase and subsequent chemiluminescence development using the ECL Western blotting system (Amersham Life Sciences, Arlington Heights, IL).

For metabolic labeling,  $10^6$  cells were labeled with 100  $\mu$ Ci of [ $^{35}$ S] methionine in 1 ml of Dulbecco's modified Eagle's medium minus methionine  
35 (Amersham Life Sciences) for 16 h. Immunoprecipitation of *RAFTK* protein from labeled cells with *RAFTK* antiserum or with normal rabbit serum (NRS) was performed as previously described (Bennett, B.D. et al. (1994) *J. Biol. Chem.* 269,

1068-1074; Yarden, Y et al. (1987) *EMBO J.* 6, 3341-3351). For immunoprecipitation of Tyr (P) proteins, cold soluble extracts were first incubated with *RAFTK* antibodies (R-4250) overnight at 4°C. The extracts were then incubated with protein-G-Sepharose beads precoupled to goat anti-rabbit IgG for 1.5 h at 4°C.

- 5 Proteins were eluted from the beads by heating the samples at 100°C for 5 min in SDS-polyacrylamide gel electrophoresis buffer. Proteins were separated by SDS-PAGE, transferred and immunoblotted with PY-20 (diluted 1:5000). The immunoreactive bands were visualized using the ECL system.

## 10 Antibodies

- Anti-*RAFTK* antiserum was obtained from New Zealand white rabbits immunized with a bacterially expressed fusion protein consisting of the GST-C-terminal (681-1009 amino acid residues) of human *RAFTK* cDNA subcloned into the pGEX-2T expression vector. The sera were titered against the GST-*RAFTK* C-
- 15 terminus fusion protein by ELISA (Dymecki, S. M. et al. (1992) *J. Biol. Chem.* 267, 4815-4823; Bennett, B.D. et al. (1991) *J. Biol. Chem.* 266, 23060-23067) and the serum (R-4250) exhibiting the highest titer (1:256,000) was used in subsequent experiments.

## 20 Isolation and characterization of *RAFTK* cDNAs

- To identify tyrosine kinases in human megakaryocytes, PCR primers based on conserved sequences of PTKs were used (Wilks, A.F. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1603-1607). RNA from the human megakaryocytic CMK cell line was used as a template to synthesize CMK cDNA. The cDNA was amplified by using the PTK
- 25 primers. Fragments of the expected size (~160 bp) were isolated and subcloned for sequence analysis. One clone that appeared to represent a novel tyrosine kinase (termed JJ3) was used as a probe to screen the human hippocampus cDNA library. A partial cDNA clone (termed S2-3) containing an ~2.0 kb insert was isolated. A homology analysis of this clone to human pp125<sup>FAK</sup> was performed and regions
- 30 were chosen to design specific primers to generate an *RAFTK* gene-specific probe. The JJ3 fragment was used to screen the human hippocampus cDNA library to obtain overlapping cDNAs. The 5'- end of each of these clones was in turn used as probes to obtain the full-length *RAFTK* cDNA. Eight different overlapping sequences were obtained of the coding region of *RAFTK*. Figure 1 is a schematic representation along
- 35 with a restriction map of the sequence showing the pattern of overlapping cDNAs. The 3.6 kb length of the *RAFTK* cDNA contains an open reading frame with the first in frame ATG codon located at nucleotide 294-296, followed by a stop codon at

position 3260-3262. This open reading frame encodes a predicted protein of 1009 amino acid residues with a calculated molecular weight of ~123 Kd and has been given the name *RAFTK* (for a related adhesion focal tyrosine kinase). Analysis of the hydrophobicity of the predicted protein revealed lack of a transmembrane region and no recognizable sites for acylation. The kinase domain is flanked by large N-terminal (426 residues) and C-terminal (331 residues) domains. Comparison of the nucleotide sequence and the deduced amino acid sequence of the encoded protein with the National Biomedical Research Foundation and GenBank data bases revealed that this cDNA encoded a tyrosine kinase related to the pp125<sup>FAK</sup>. The predicted amino acid sequence of pp120<sup>RAFTK</sup> contains the structural motifs common to all protein kinases, including the putative ATP-binding site (432-437a.a. Gly<sup>432</sup> - Xaa - Gly<sup>434</sup> - Xaa - Xaa - Gly<sup>437</sup>), and three residues that are predicted to interact with the  $\gamma$ -phosphate group of the bound ATP molecule (in positions 402a.a, 529a.a and 655a.a). In addition, *RAFTK* contains two peptide sequences that are highly conserved among PTKs (Asp<sup>549</sup> - Ile<sup>550</sup> - Ala<sup>551</sup> - Val<sup>552</sup> - Arg<sup>553</sup> - Asn<sup>554</sup> and Pro<sup>588</sup> - Ile<sup>589</sup> - Lys<sup>590</sup> - Trp<sup>591</sup> - Met<sup>592</sup>). Interestingly, like chicken pp125<sup>FAK</sup>, the C-terminal region of *RAFTK* contains a proline-rich stretch (residues 690-767) where the proline content exceeds 20%. A unique domain is found at the NH<sub>2</sub> terminus of *RAFTK* (amino acids 1-39) (Figure 3). This region is the most divergent among various PTKs and may be involved in cellular localization and/or interaction with other cellular proteins. Like pp125<sup>FAK</sup>, *RAFTK* does not contain SH2 or SH3 domains. The kinase domain (amino acid 427 - 679) of *RAFTK* shares 60% identical homology with the mouse pp125<sup>FAK</sup>, 54% with human pp125<sup>FAK</sup>, and 36% with src (Figure 2). The kinase domain consists primarily of the catalytic domain including the putative ATP binding site (amino acids 432 - 437). *RAFTK* shares 42% homology in the NH<sub>2</sub> domain and about 39% in the C-terminal domain with mouse pp125<sup>FAK</sup>. The overall amino acid homology of *RAFTK* is 48% identity (65% similarity) with mouse pp125<sup>FAK</sup>.

### 30 **Molecular cloning of the full-length murine *RAFTK* cDNA**

Southern blot analysis of human and mouse genomic DNA digested with EcoRI, HindIII, BamHI, XbaI, PstI and probed under conditions of high stringency with 3'- fragment of *RAFTK* cDNA from 1595 - 2974 bp (1.4 kb) as a probe, revealed a single band in each lane, indicating that the human *RAFTK* gene and the mouse *RAFTK* gene are highly homologous and are single genes. Therefore, a random and oligo (dT)-primed mouse adult brain cDNA library was screened under conditions of high stringency for the full-length mouse cDNA of *RAFTK* using the 5'- fragment and

3'- fragment of human *RAFTK* cDNA as probes. Four clones were isolated and two of these clones were sequenced in both directions and additional clones were partially sequenced. Sequence analysis of these clones revealed identical sequences. The 4.5 kb full length cDNA has an open reading frame of 1009 amino acid residues and possesses 95.6% identical homology with the human *RAFTK* gene.

#### Chromosomal localization of human *RAFTK* gene

Hamster, human and mouse DNAs were digested with BamHI, HindIII and PstI to identify a specific RFLP pattern for the *RAFTK* gene in each species. Southern blots were probed with a human 1.4 kb *RAFTK* cDNA. Unique BamHI 16.5 kb and 14.5 kb fragments for *RAFTK* were identified in human DNA from the parental cell lines used to prepare human/rodent cell hybrids. DNAs from the parental and the somatic hybrid cell lines in mapping panel #2 were digested with BamHI, Southern blotted and probed. Analysis indicated that the human-specific BamHI pattern was observed in cell line #8 which contains human Chromosome 8 (Figure 4). A fainter signal was also observed for the human-specific BamHI pattern in hybrid cell line #20 (Fig. 4) which, although it contained an intact human Chromosome #20, also carried a gene from human Chromosome #8 (NEFL, neurofilament light polypeptide, 8p21) as determined by Southern blot hybridization (Coriel Cell Institute for Medical Research, Camden, NJ). All other hybrid cell lines were negative for the human-specific BamHI RFLP. Additionally, when the human 1.4 kb *RAFTK* cDNA was used to probe Coriel Panel #1, the human-specific fragment was detected in all hybrids containing greater than 4% of human Chromosome 8 and was absent in every hybrid that lacked Chromosome 8.

Southern blots of C57BL/6J and Mus spretus DNAs were digested with 29 different restriction enzymes and probed with a human *RAFTK* 1.4 kb cDNA. A BamHI restriction fragment length polymorphism (RFLP) was detected. The alleles for this BamHI RFLP consist of 8.6 kb and 5.2 kb genomic DNA bands, characteristic of C57BL/6J, and 15.5 kb and 6.7 bands which are found in Mus spretus. These alleles were characterized in 87 DNAs from the C57BL/6J X Mus spretus backcross panel. Results of the haplotype analysis from this mapping data indicate that the *RAFTK* gene co-localizes with D14Bir10 (DNA segment-Birkenmeier 10) and is linked to Nfl (neurofilament, light polypeptide) on mouse Chromosome 14 (Figure 5). The *RAFTK* locus mapped between Xmv19 (xenotropic-MCF leukemia virus-19) and Nfl and the calculated map distances for these loci are: Xmv19,  $7.1 \pm 5.3$  cM, *RAFTK*,  $3.5 \pm 2.0$  cM, Nfl.

The position of *RAFTK* on mouse Chromosome 14 was confirmed by determining the segregation of a *SacI* RFLP for *RAFTK* DNAs from BXD recombinant inbred (RI) lines. The *SacI* RFLP for *RAFTK* was indicated by the presence of a 16.5 kb genomic DNA band in C57BL/6J or a 6.2 kb fragment in DBA/2J. These alleles were characterized for 26 DNAs from the BXD RI line. The strain distribution patterns of *RAFTK* and the locus coding for gonadotropin releasing hormone, *Gnrh* (Hearne, C.M. et al. (1991) *Mamm. Genome* 1, 273-282), indicate close linkage between these two loci on Chromosome 14. Perfect concordance was observed with the BXD strain distribution pattern for the *Gnrh* locus, indicating linkage of less than one map unit distance from *RAFTK* *Gnrh* (Silver, J. (1985) *J. Hered.* 76, 436-440). This mapping data places *RAFTK* distal to *Nfl* and is a contradiction to the backcross data. However, backcross data are not as accurate as RI data since backcross mice were derived from interspecies cross.

#### 15 **Expression of *RAFTK* in tissues and cell lines**

A specific *RAFTK* probe was designed (nucleotide 2958 bp - 3103 bp). This sequence is present in *RAFTK* and not in human pp125<sup>FAK</sup>. This probe was used for hybridization of all Northern blots described here.

Northern blot analysis of RNA from human fetal heart, brain, lung, liver and kidney revealed a weak single major species of mRNA of 4.5 kb in brain and it appears to be expressed at low levels in the lung and liver. Expression in human adult tissues was assessed by hybridization of the cDNA probe to a Northern blot of poly (A<sup>+</sup>) RNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. While heart and skeletal muscle RNA samples were negative for *RAFTK*, a single mRNA was observed in all other tissues with the highest levels expressed in brain. To further characterize the distribution of *RAFTK* expression in other human tissues, Northern blot analysis of spleen, thymus, prostate, testes, ovary, intestine, colon and peripheral blood leukocytes revealed high expression of *RAFTK* in thymus, spleen and peripheral blood leukocytes. Northern blot analysis of different human brain regions (amygdala, caudate nucleus, corpus callosum, hippocampus, hypothalamus, substantia nigra, subthalamic nucleus and thalamus) revealed that the highest expression of *RAFTK* was in amygdala and hippocampus. Lower expression was observed in the other brain regions, with the exception of corpus callosum and substantia nigra where there was no detectable signal. These results demonstrate that brain has abundant expression of *RAFTK*, especially in the amygdala and hippocampus.

Expression of *RAFTK* was observed in several megakaryocytic cell lines such as CMK, Mo7c, HEL and DAMI cells. In addition, expression of *RAFTK* was detected in Ramos, FHS and HeLa cells but low level of expression was detected in Jurkat, Hep 3B and CCL 75 cells. Using PCR techniques, expression of *RAFTK* was also found in primary bone marrow megakaryocytes, blood platelets, and in marrow CD34<sup>+</sup> progenitor cells. The level of expression of *RAFTK* mRNA is similar to FAK in CD34<sup>+</sup> cells, and is higher than FAK in bone marrow megakaryocytes. In platelets, the level of expression of *RAFTK* mRNA is lower than FAK, as observed by PCR under the same experimental conditions. *RAFTK* mRNA expression in bone marrow megakaryocytes is higher than that in CD34<sup>+</sup> cells. Taken together, these results demonstrate that *RAFTK* is abundantly expressed in brain and hematopoietic cells. The restricted expression observed in fetal versus adult tissues indicates its expression is upregulated during development.

#### 15 **Generation of specific antibodies for *RAFTK* and detection of *RAFTK* protein**

The fusion protein GST-C-terminus of *RAFTK* (residues 681-1009) was chosen for rabbit immunizations in order to obtain specific antibodies for *RAFTK* protein. These polyclonal antibodies (R-4250) do not cross react with pp125<sup>FAK</sup>. The monoclonal antibody 2A7 against FAK does not cross react with the C-terminal GST-*RAFTK* fusion protein, indicating that *RAFTK* might be antigenically different from FAK. Furthermore, FAK immunoprecipitated by the monoclonal antibody 2A7 from megakaryocytes was not recognized by polyclonal antiserum 4250. Similarly, *RAFTK* immunoprecipitated by antiserum 4250 also was not recognized by the monoclonal antibody 2A7. Taken together, these data demonstrate that FAK and *RAFTK* are distinguishable antigenically while being related members of the FAK family.

The specificity of this antiserum was examined by immunoprecipitation. The CMK cell line was metabolically labeled with [<sup>35</sup>S] methionine, and extracts were immunoprecipitated with anti-*RAFTK* antiserum. A major protein species of ~123 Kd was detected in CMK cells. A similar species was observed in other human megakaryocytic cell lines such as DAMI. This band was not observed when normal rabbit serum or pre-immune rabbit serum was used for immunoprecipitation. Incubation of R4250 with 1 µg or 10 µg of the C-terminus of GST-*RAFTK* fusion protein abolished the appearance of ~123 Kd, while incubation with 10 µg of the fusion protein GST-MATK-SH2 domain did not have any effects. These results demonstrate that polyclonal antibodies R-4250 are specifically recognizing *RAFTK* protein of ~123Kd size. Furthermore, thrombin (1 unit/ml) stimulated a rapid increase

in the amount of *RAFTK* protein immunoreactivity in anti-Tyr(P) immunoprecipitates. These results demonstrate that *RAFTK* is a protein tyrosine kinase, and that thrombin can induce its tyrosine phosphorylation.

The method of PCR cloning has been successfully employed by many laboratories to identify novel members of the PTK family. Using this strategy, a novel intracytoplasmic tyrosine kinase in human megakaryocytic cells has been identified, termed *RAFTK*. Sequence analysis of *RAFTK* revealed ~48% identity (65% similarity) to pp125<sup>FAK</sup> suggesting that *RAFTK* belongs to this subfamily of cytoplasmic tyrosine kinases. *RAFTK* does not appear to be the recently described FAKB protein (Kanner, S.B. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91, 10484-10487), also related to pp125<sup>FAK</sup>, since the specific amino acid sequence used to make antisera which recognized the FAKB protein is missing in the predicted amino acid sequence of *RAFTK* protein. Furthermore, unlike FAKB, *RAFTK* protein did not form stable complexes with the TCR/CD3 linked tyrosine kinase ZAP 70 in T-cells indicating that *RAFTK* and FAKB are different proteins.

The chicken, human and mouse focal adhesion kinases have been recently implicated as playing key roles in signal transduction pathways associated with extracellular adhesion molecules and with receptors for neuropeptide growth factors (Schaller, M.D., and Parsons, J.T. (1993) *Trends Cell Biol.* 3, 258-262; Zachary, I., and Rozengurt, E. (1992) *Cell* 71, 891-894; Leeb-Lundberg, L.M., and Song, X.-H. (1991) *J. Biol. Chem.* 266, 7746-7749; Zachary, I. et al. (1992) *J. Biol. Chem.* 267, 19031-19034). Thus, based on its homology to pp125<sup>FAK</sup>, one would expect *RAFTK* to participate in signalling pathways as well. The deduced 1009 amino acid sequence of *RAFTK* (with calculated molecular mass of 120 Kd) contains a kinase domain and lacks a transmembrane region, myristylation sites, and SH2 and SH3 domains. In order to identify conserved regions within *RAFTK* between species that may have important functions, the murine homolog of the human *RAFTK* cDNA was cloned. The sequence identity between the human and murine *RAFTK* cDNAs is 90% at the nucleotide level and 95.6% at the predicted amino acid level. In the kinase domain, 98.5% of the amino acids are identical. Therefore, the *RAFTK* gene is highly conserved in human and rodent, again suggesting an important role in cell signalling functions. The *RAFTK* has an insertion of an additional 4 amino acids between 76-81 (G<sup>76</sup>R<sup>77</sup>I<sup>78</sup>G<sup>79</sup>) compared to chicken, murine, and human pp125<sup>FAK</sup> sequences (Schaller, M.D. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196; Schaller, M.D., and Parsons, J.T. (1993) *Trends Cell Biol.* 3, 258-262; Clark, E.A., and Brugge, J.S. (1995) *Science* 268, 233-239). Amino acids corresponding to positions 292 - 320 of human pp125<sup>FAK</sup> and amino acids corresponding to position 850 - 864 and 901 -



926 of chicken pp125<sup>FAK</sup> are absent in the predicted *RAFTK* protein. Interestingly, like chicken pp125<sup>FAK</sup>, the C-terminus region of human *RAFTK* and mouse *RAFTK* contains a proline-rich stretch (residues 690-767). It has been shown that proteins containing proline-rich peptide motifs (such as Shc, p62 and ribonucleoprotein K) could serve as SH3 domain ligands, and that the binding of these proteins to the Src SH3 domain was inhibited with a proline-rich peptide ligand (Weng, Z. et al. (1994) *Mol. Cell. Biol.* 14, 4509-4521). Furthermore, the predicted *RAFTK* protein, like the pp125<sup>FAK</sup> protein, displays several unique features among the known tyrosine kinases. The primary sequence of *RAFTK* does not contain a signal peptide or a membrane-spanning region and the protein is therefore presumed to be located in the cytoplasm. *RAFTK* lacks SH2 and SH3 domains, which are structural elements involved in protein-protein interactions (Pawson, T., and Gish, G.D. (1992) *Cell* 71, 359-362; Konopka, J. B. et al. (1984) *J. Virol.* 51, 223-232; Waksman, G. et al. (1992) *Nature* 358, 646-653; Taylor, S.J., and Shalloway, D. (1993) *Current Opin. Genet. & Dev.* 3, 26-34; Yu, H. et al. (1992) *Science* 258, 1665-1668), and does not exhibit significant homology with any known PTK beyond pp125<sup>FAK</sup> outside of the catalytic domain. Lack of SH2 and SH3 domains suggests that other regions within *RAFTK* protein are important for protein interaction and targeting. In the case of the pp125<sup>FAK</sup> protein, it has been demonstrated by structural-functional analysis that 159 amino acids within the C-terminus are essential as a "Focal adhesion targeting" sequence (Hildebrand, J.D. et al. (1993) *J. Cell. Biol.* 123, 993-1005). The homology between *RAFTK* and pp125<sup>FAK</sup> within this region is 52%. The overall structure of *RAFTK* is characteristic of the pp125<sup>FAK</sup> gene, with the catalytic domain flanked by large N-terminal and C-terminal domains. It has recently been reported that deletions of the NH<sub>2</sub> - or the COOH-terminal non-catalytic domain of pp125<sup>FAK</sup> including Tyrosine<sup>397</sup> did not abolish the kinase activity of pp125<sup>FAK</sup> (Chan, P.-Y. et al. (1994) *J. Biol. Chem.* 269, 20567-20574). Moreover, there is conservation of several tyrosine residues between *RAFTK* and pp125<sup>FAK</sup> including Tyrosine<sup>397</sup> which has been shown to be the major site of tyrosine phosphorylation in pp125<sup>FAK</sup> protein (Schaller, M.D. et al. (1994) *Mol. Cell. Biol.* 14, 1680-1688).

*RAFTK* specific mRNA expression was observed in human fetal tissues most abundant in brain (predominantly in amygdala and hippocampus regions) and appeared to be developmentally upregulated as demonstrated in the pattern of adult tissue expression. Within the hematopoietic system, in addition to peripheral blood leukocytes, a high level of specific mRNA expression of *RAFTK* was detected in B-

cells and various megakaryocytic cell lines. By using PCR, the specific mRNA expression of *RAFTK* was also detected in primary bone marrow CD34<sup>+</sup> progenitor cells, primary bone marrow megakaryocytes and platelets.

- RAFTK* is phosphorylated after thrombin treatment of CMK cells. FAK protein was also found phosphorylated on tyrosine after thrombin or collagen treatment of platelets (Lipfert, L. et al. (1992) *J. Cell Biol.* 119, 905-912). There is considerable homology in the thrombin receptors and considerable signal similarities in transduction mechanisms between platelets and megakaryocytes (Vittet, D., and Chevillard, C. (1993) *Blood Coagulation & Fibrinolysis* 4, 759-768). Furthermore, bone marrow megakaryocytes in liquid culture stimulated with thrombin for 5 min revealed dramatic morphological changes reminiscent of those found in platelets, including shape change and organelle centralization that involved immature as well as mature cells (Cramer, E.M. et al. (1993) *Am. J. Path.* 143, 1498-1508). Megakaryocytes were also able to secrete alpha-granule proteins in the dilated cisternae of the demarcation membrane system (Cramer, E.M. et al. (1993) *Am. J. Path.* 143, 1498-1508).

- The human *RAFTK* gene was found on Chromosome 8 using DNAs from the somatic cell hybrid lines. The signal observed in cell line #20 in mapping panel #2 suggested that a fragment of Chromosome 8 is in the Chromosome #20 cell line. Although cell line #20 contained the human NEFL gene, there was no evidence for Chromosome 20 or a fragment of Chromosome 20 in cell line #8 (Coriel Cell Institute for Medical Research, Camden, NJ). The localization of *RAFTK* to Chromosome 8 was confirmed using mapping panel #1. The human NEFL gene has been localized to Chromosome 8p21 (Hurst, J. et al. (1987) *Cytogenet. Cell Genet.* 45, 30-32). Nfl, the murine homolog of human NEFL, has been mapped to mouse Chromosome 14 and is within 3 cM of the Gnrh locus (GBASE). The close linkage of the mouse *RAFTK* gene to Nfl (whose homolog NEFL is on human Chromosome 8p21) suggested that the human *RAFTK* gene may be mapped to Chromosome 8 based on homology between human and mouse chromosomes (Hurst, J. et al. (1987) *Cytogenet. Cell Genet.* 45, 30-32). Therefore, the human *RAFTK* gene is localized to Chromosome 8p21. The mouse *RAFTK* gene has been mapped to Chromosome 14 using a (C57BL/6J) x M. spretus F<sub>1</sub> x M. spretus backcross. The position of mouse *RAFTK* was confirmed by RI line mapping using the BXD RI lines. The *RAFTK* gene was also shown to be closely linked to Gnrh whose human homolog (LHRH-luteinizing hormone releasing hormone) has been mapped to human Chromosome 8p21-11.2 (Yang-Feng, T.L. et al. (1986) *Somatic Cell. Mol. Genet.* 12, 95-100).

**EXAMPLE 2:           Activation of the Novel Protein Tyrosine Kinase, *RAFTK*,  
in Megakaryocytes Upon SCF and PMA Stimulation and  
Its Direct Association with Paxillin**

5           *RAFTK* appears to be a member of the Focal Adhesion Kinase (FAK) family, and is involved in  $\text{Ca}^{2+}$ -mediated signalling events in PC-12 cells. In this Example, the signalling pathways involving *RAFTK* in human megakaryocytic cells were characterized. Stem Cell Factor (SCF), which potentiates the growth of megakaryocytes and their progenitors, and Phorbol Myristate Acetate (PMA), which  
10 causes differentiation of megakaryocytic cell lines, induced the tyrosine phosphorylation of *RAFTK* through Protein Kinase C (PKC). The constitutive association of *RAFTK* with PKC- $\delta$  was observed, while the association of *RAFTK* with PKC- $\alpha$  was induced upon stimulation with SCF. In addition, the direct association of *RAFTK* with paxillin, a 68-Kd cytoskeleton protein, was demonstrated.  
15 Upon the activation of *RAFTK*, there was a sequential activation and phosphorylation of paxillin. Cytochalasin D, which disrupts the cytoskeleton, abolished the phosphorylation of *RAFTK* upon PMA and SCF stimulation.

          These results show that *RAFTK* is a down-stream signalling protein of PKC and that paxillin is a down-stream associated protein of *RAFTK*. Furthermore,  
20 *RAFTK* association with the cytoskeleton was critical for its phosphorylation. These observations show the manner in which *RAFTK* participates in megakaryocyte proliferation and differentiation.

          The following materials and methods were used to study activation of *RAFTK*:  
25

**Materials**

          Recombinant SCF/KL and polyclonal anti-c-kit antibodies were generously provided by Dr. Keith E. Langley and Dr. L. Bennett, Amgen Inc. (Thousand Oaks, CA). Monoclonal anti-phosphotyrosine antibody (PY-20) and monoclonal anti-  
30 paxillin were obtained from ICN (Costa Mesa, CA ); monoclonal antibodies anti-p85, anti-Shc, anti-Grb2, anti-FAK, anti-PKC- $\alpha$ , anti-PKC- $\beta$ , and anti-PKC- $\delta$  were obtained from Transduction Laboratories (Lexington, KY). Calphostin C, staurosporine, calcium ionophore A23187, EGTA and Phorbol 12-Myristate 13-Acetate (PMA) were obtained from Calbiochem (La Jolla, CA ). Electrophoresis  
35 reagents were obtained from Bio-Rad Laboratories (Hercules, CA). All other reagents were purchased from Sigma Co. (St. Louis, MO).

### CMK cells

The CMK cell line, provided by Dr. T. Sato (Chiba University, Japan), was maintained in RPMI 1640 with 10% fetal calf serum (FCS) as described previously (Sato, T. et al. (1989) *Br. J. Hematol.* 72, 184-190). The CMK cell line was derived from a child with megakaryoblastic leukemia and has properties of cells of the megakaryocytic lineage, including the surface expression of glycoproteins Ib and IIb/IIIa, synthesis of platelet factor 4, PDGF and von Willebrand factor. CMK cells can proliferate in response to cytokines and have been used by us and other investigators in studies of megakaryocyte growth and maturation. In addition, CMK cells also differentiate upon induction with PMA (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-79; Namciu, S. et al. (1994) *Oncogen* 9, 1407-1416). For such experiments, PMA was dissolved in dimethyl sulfoxide and stored at -20°C until use, when it was diluted in RPMI 1640 medium.

### Antibodies

Anti-*RAFTK* antiserum was obtained from New Zealand White rabbits immunized with a bacterially expressed fusion protein consisting of GST and the C-terminus (amino acids 681-1009) of human *RAFTK* cDNA subcloned into the pGEX-2T expression vector as described (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10). The sera were titred against the GST-*RAFTK* C-terminus fusion protein by an enzyme-linked immunosorbent assay, and the serum (R-4250) exhibiting the highest titer (1:256,000) was used in subsequent experiments. In immunoblotting experiments anti-phosphotyrosine antibodies were used (PY20) or anti-*RAFTK* (1:1000) or antibodies for FAK-2A7 (1:1000).

### Cell stimulation, immunoprecipitation and immunoblotting

The CMK cells were starved overnight in RPMI-1640 with 0.5% FCS. Cells (106/ml) were stimulated for 0 to 30 min at RT with either SCF (100 to 500 ng/ml) or PMA (10-100 nM). The stimulation was terminated by adding ice-cold RPMI-1640 containing sodium vanadate followed by centrifugation. The cells were lysed in modified-RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 2 (g/ml of aprotinin, leupeptin and pepstatin, and 1 mM Na<sub>3</sub>VO<sub>4</sub>). Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by using a protein assay (Bio-Rad Laboratories) and were standardized to equal concentrations of protein prior to immunoprecipitation. Identical amounts of protein

from each sample were precleared by incubation with Protein G-Sepharose CL-4B (Sigma Co., St. Louis, MO) for 1 h at 4°C. After the removal of Protein G-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as described below for each experiment for 4 h or overnight at 4°C.

- 5 Immunoprecipitation of the antigen-antibody complex was accomplished by incubation for 1 h at 4°C with 40 µl of protein G-Sepharose as described (Huang, E. et al. (1990) *Cell* 63, 225-233). Normal rabbit serum was used as a control in immunoprecipitations. Bound proteins were solubilized in 20 µl of 2 X Laemmli buffer. Samples were separated and analyzed by 7.5% SDS-PAGE, and then
- 10 transferred to nitrocellulose membranes. The membranes were blocked with 5% bovine serum albumin (BSA) (Boehringer Mannheim Co., Indianapolis, IN) and probed with primary antibody for 1 h at RT. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) reagents (Amersham Corp., Arlington Heights, IL).

15

**SCF induces the tyrosine phosphorylation of *RAFTK* in CMK megakaryocytic cells.**

- To investigate whether *RAFTK* participated in growth-related signal transduction pathways in megakaryocytes, model CMK megakaryocytic cells were
- 20 studied with a focus on the c-kit/SCF pathway which is known to be important in the proliferation of this lineage (Briddell, R. A. et al. (1991) *Blood* 78, 2854-2859; Avraham, H. et al. (1992) *Blood* 79, 365-371; Avraham, H. et al. (1992) *Blood* 80, 1679-1684). The CMK cells were starved in RPMI-1640 medium followed by stimulation with SCF, and then harvested at different time intervals as indicated. Cells
- 25 were lysed and immunoprecipitated with polyclonal *RAFTK* R-4250 antibodies. The precipitates were then immunoblotted with monoclonal anti-phosphotyrosine (PY-20) antibodies or FAK monoclonal antibodies or with *RAFTK*-specific antibodies.

- Tyrosine phosphorylation of *RAFTK* peaked at a concentration of SCF of 500 ng/ml at 5 min. No phosphorylation of FAK was observed under these conditions.
- 30 Maximum stimulation of *RAFTK* in CMK cells was observed within 1 min and peaked at 5 min.

- SCF treatment is known to increase cytoplasmic calcium levels (Columbo, M. et al. (1994) *Biochemical Pharmacology* 47, 2137-2145) as well as to stimulate phosphorylation of components of c-kit associated signalling pathways (Lev, S. et al.
- 35 (1992) *Journal of Biological Chemistry* 267, 15970-15977; Rottapel, R. et al. (1991) *Molecular & Cellular Biology* 11, 3043-3051). Since SCF effects are modulated by PKC (Blume-Jensen, P. et al. (1995) *Journal of Biological Chemistry* 270, 14192-

14200; Namciu, S. et al. (1994) *Oncogene* 9, 1407-1416; Grabarek, J. et al. (1992) *Journal of Biological Chemistry* 267, 10011-10017), the role of PKC in *RAFTK* stimulation was investigated. SCF treatment of CMK cells induced rapid phosphorylation of *RAFTK* within 1 min and was completely blocked by the PKC inhibitors calphostin C or staurosporine.

**PMA induces tyrosine phosphorylation of *RAFTK* in CMK megakaryotic cells.**

To determine if *RAFTK* participates in pathways of megakaryocyte differentiation, the effect of PMA, which induces differentiation and maturation of cells of this lineage (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-79; Namciu, S. et al. (1994) *Oncogene* 9, 1407-1416), was studied and was found to be able to stimulate *RAFTK* phosphorylation. Under these conditions, phosphorylation of *RAFTK* peaked at 5 min at 50 nM PMA.

Because PMA activates PKC, the relationship between PMA induced tyrosine phosphorylation of *RAFTK* in CMK cells and PKC activation was further characterized. Addition of the PKC inhibitors Calphostin C (1  $\mu$ M) or staurosporine (75 nM) inhibited *RAFTK* phosphorylation following PMA treatment of CMK cells. In additional correlative experiments, PMA-sensitive isozymes were first down-regulated in CMK cells by prolonged treatment with PMA (15 min at 37°C), and then these cells were treated with SCF or PMA. This prolonged treatment with PMA completely abolished the subsequent effects of PMA or SCF treatment, suggesting that tyrosine phosphorylation of *RAFTK* by SCF or PMA is a PKC dependent mechanism. PMA stimulation did not result in an increase in the tyrosine phosphorylation of FAK, suggesting these related molecules may have distinct roles in different signalling activation pathways in megakaryocytes.

**Shc association with Grb2 in SCF stimulated CMK cells.**

The activated c-kit receptor can recruit Grb2 by tyrosine phosphorylation of Shc (Liu, L. et al. (1994) *Mol. and Cell Biol.* 14, 6926-6935; Tauchi, T. et al. (1994) *J. Exp. Med.* 179, 167-175). It was then investigated whether SCF could induce tyrosine phosphorylation of Shc and whether it associates with Grb2 in CMK cells. CMK cells stimulated by SCF (500 ng/ml) or PMA (50 nM) were lysed and immunoprecipitated with monoclonal anti-human Shc antibody and the precipitates were then immunoblotted with monoclonal anti-phosphotyrosine antibody PY-20. SCF induced Shc phosphorylation and association with Grb2 while PMA has no

effect on Shc-Grb2 association, suggesting differences in proliferative versus maturational pathways with regard to these molecules in megakaryocytes.

**PKC $\alpha$  and PKC $\delta$  isoforms are involved in *RAFTK* phosphorylation.**

5 To identify the PKC-isozymes involved in *RAFTK* phosphorylation, an analysis of PKC isozymes present in CMK cells was performed. PKC $\delta$  and PKC $\alpha$  isoforms were expressed in megakaryocytes as observed by immunoprecipitation using specific antibodies for PKC $\alpha$  and PKC $\delta$ , while no expression of PKC $\beta$  or PKC $\gamma$  was observed. To characterize which PKC-isozymes that may be involved in *RAFTK* phosphorylation upon SCF or PMA stimulation of CMK cells, CMK cell lysates  
10 unstimulated or after PMA or SCF treatment were immunoprecipitated with either PKC $\delta$  or PKC $\alpha$  specific antibodies. The immunoprecipitates were resolved on 7.5% SDS-PAGE, immunoblotted with PY-20 antibodies or anti-phosphoserine/threonine or PKC $\delta$  or PKC $\alpha$  or *RAFTK* specific antibodies. Constitutive association of PKC $\delta$   
15 with *RAFTK* was demonstrated by co-immunoprecipitation studies as shown in. An increase in phosphorylation of PKC $\delta$  upon PMA treatment was observed. A similar observation of augmented PKC $\delta$  phosphorylation was made in CMK cells stimulated with SCF.

PKC $\alpha$  was found only to be associated with *RAFTK* after stimulation with  
20 SCF or PMA, but changes in the degree of phosphorylation of PKC $\alpha$  upon such activation by SCF or PMA using phosphoserine/threonine antibodies were not detected. These results show that PKC $\alpha$  and PKC $\delta$  isozymes are involved in *RAFTK* activation, that PKC $\delta$  is constitutively associated with *RAFTK* while association of *RAFTK* with PKC $\alpha$  is inducible, and may be increased in activation upon certain  
25 stimuli but may not alter the phosphorylation status of this species.

**Association of *RAFTK* with Paxillin and PKC $\delta$ .**

To address the role of *RAFTK* in the formation of focal adhesions, *RAFTK* association with a known focal adhesion protein paxillin was investigated. Cell  
30 lysates of CMK cells treated with PMA were immunoprecipitated with either *RAFTK* specific antibodies or anti-paxillin antibodies. The immunoprecipitates were resolved and immunoblotted with PY-20 antibodies. An increase in phosphorylation of paxillin was observed upon PMA stimulation which peaked by 5 min. Constitutive association of paxillin and *RAFTK* was observed in untreated CMK cells.

35 Treatment with the PKC inhibitor calphostin C abolished *RAFTK* activation and decreased its association with paxillin. Paxillin phosphorylation at 10 min was not altered by calphostin C treatment, indicating that paxillin phosphorylation is not

dependent on PKC. Similar observations of paxillin phosphorylation and its association with *RAFTK* were observed in CMK cells stimulated with SCF.

In addition, CMK cells stimulated with SCF were immunoprecipitated with *RAFTK* antibodies, and the immunoprecipitates were then resolved and  
5 immunoblotted with PY-20 antibodies, *RAFTK* or paxillin antibodies. *RAFTK* constitutively associated with paxillin and was not altered in its degree of phosphorylation upon activation with SCF. Constitutive association of PKC $\delta$  with *RAFTK* was also observed. These results demonstrate a constitutive association of PKC $\delta$ , *RAFTK* and paxillin in untreated cells. Upon activation with PMA, PKC $\delta$  was  
10 phosphorylated and peaked by 10 min. The constitutive association of PKC $\delta$  with *RAFTK* was not affected by PMA stimulation.

***RAFTK* activation is inhibited by BAPTA, calphostin C and cytochalasin-D.**

Calcium ionophore (A23187) treatment of cells elevates intracellular calcium  
15 levels and initiates a cascade of signalling events including PKC activation. CMK cells treated with calcium ionophore A23187 showed tyrosine phosphorylation of *RAFTK*, which was inhibited by the intracellular calcium chelator, BAPTA. In the presence of calphostin C, a specific PKC inhibitor, induction of *RAFTK* phosphorylation by the calcium ionophore A23187 was inhibited, indicating that  
20 calcium regulation of this PKC isoform was involved in *RAFTK* phosphorylation. Upon calcium ionophore A23187 treatment, *RAFTK* activation was completely inhibited in the presence of cytochalasin-D, indicating that *RAFTK* is associated with the cytoskeleton and this association is essential for its activation following changes in intracellular calcium.

25 Similarly, SCF treatment of CMK cells induced rapid, transient tyrosine phosphorylation of *RAFTK* which was inhibited in the presence of BAPTA, suggesting that SCF may induce *RAFTK* phosphorylation through elevating intracellular calcium levels. However, in the presence of cytochalasin-D, SCF induction of *RAFTK* phosphorylation was inhibited completely, indicating again that  
30 the integrity of the cytoskeleton is required for *RAFTK* phosphorylation. Since SCF stimulation of *RAFTK* in the presence of calphostin C was also completely inhibited, this mechanism of *RAFTK* activation appears to be mediated through PKC.

PMA phosphorylation of *RAFTK* also was blocked by calphostin C or by BAPTA, further indicating that calcium regulation of PKC isoforms is involved in  
35 *RAFTK* stimulation. Cytochalasin-D treatment inhibited PMA stimulation of *RAFTK*, suggesting that *RAFTK* association with the megakaryocytic cytoskeleton is critical for its phosphorylation in cells of this lineage.



In this study, *RAFTK* activation and its regulation in megakaryocytic cells was characterized. The results obtained in these studies demonstrated that *RAFTK*, unlike FAK, is tyrosine phosphorylated upon SCF and PMA treatments. The finding that FAK is not phosphorylated under these conditions is consistent with prior studies of FAK activation in Mo7E megakaryocytic cells (Gotoh, A. et al. (1995) *Experimental Hematology* 23, 1153-1159) and suggests important differences in the roles of FAK and *RAFTK* in cells of this lineage.

The effects of SCF, PMA and  $\text{Ca}^{2+}$  on activation of *RAFTK* were mediated through PKC. Moreover, direct association of *RAFTK* with paxillin was observed and activation of *RAFTK* resulted in a sequential activation and phosphorylation of this cytoskeletal protein.

PKC plays an important role in cellular responses to various hormones, growth factors, neurotransmitters and cytokines, and transduces signals promoting lipid hydrolysis (See, e.g., Dekker, L. V. and Parker, P. J. (1994) *Trends in Biochemical Sciences* 19, 73-77; Nishizuka, Y. (1992) *Science* 258, 607-614; Nishizuka, Y. (1986) *Science* 233, 305-312). PKC regulates the action of a variety of ion channels, G-protein coupled receptors, tyrosine kinase receptors or non-receptor tyrosine kinases (See, e.g., Ohtani, K. et al. (1995) *Journal of Neurochemistry* 65, 605-614; Rozengurt, E. (1995) *Cancer Surveys* 24, 81-96; Sadoshima, J., et al. (1995) *Circulation Research* 76, 1-15). Since the proliferative effects of SCF appeared to be modulated in part by PKC (See, e.g., Blume-Jensen, P. et al. (1995) *Journal of Biological Chemistry* 270, 14192-14200; Sato, T. et al. (1989) *Br. J. Hematol.* 72, 184-190; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231), a role for PKC in *RAFTK* stimulation was investigated. Indeed, SCF induced rapid phosphorylation of *RAFTK* and was completely blocked by the PKC inhibitors calphostin C or Staurosporine. Furthermore, PMA, which induces differentiation of CMK megakaryocytic cells (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-79; Namciu, S. et al. (1994) *Oncogen* 9, 1407-1416), also stimulated *RAFTK* tyrosine phosphorylation; this *RAFTK* phosphorylation was abolished by the PKC inhibitors or prolonged treatment with PMA. These results also demonstrate that *RAFTK* activation was mediated by PKC.

To further characterize this pathway, the role of PKC isoforms in *RAFTK* activation was studied. The PKC- $\alpha$  and PKC- $\delta$  isoforms are known to be expressed in megakaryocytes while PKC- $\gamma$  and PKC- $\beta$  have not been found (Grabarek, J. et al. (1992) *Journal of Biological Chemistry* 267, 10011-10017). Constitutive association *in vivo* of PKC- $\delta$  with *RAFTK* was observed while association of PKC- $\alpha$  and *RAFTK*

was inducible. No increase in the level of PKC- $\alpha$  or PKC- $\delta$  with *RAFTK* after SCF or PMA stimulation was observed by co-immunoprecipitation techniques. Although PKC isoenzymes do not possess intrinsic tyrosine kinase activity, activation of PKC by phorbol esters such as PMA has been demonstrated to indirectly induce tyrosine phosphorylation in different cells types (See, e.g., Li, W. et al. (1994) *Journal of Biological Chemistry* 269, 2349-2352; Einspahr, K. J. et al. (1990) *Journal of Immunology* 145, 1490-1497; Nel, A. E et al. (1990) *Journal of Immunology* 145, 971-979). In NIH-3T3 or 32D transfectants overexpressing various PKC isoenzymes, pronounced phorbol diester-dependent tyrosine phosphorylation of PKC- $\delta$  was observed, while no detectable tyrosine specific phosphorylation was found after treatment with the other PKC isoenzymes transfectants (Li, W. et al. (1994) *Journal of Biological Chemistry* 269, 2349-2352).

SCF induction of tyrosine phosphorylation of *RAFTK* could be mediated by elevated intracellular calcium levels and activation of PKC through PLC- $\gamma$  (Yeo, E. J. et al. (1994) *Journal of Biological Chemistry* 269, 27823-27826; Ma, Y. H. et al. (1994) *Journal of Biological Chemistry* 269, 30734-30739; Zirrgiebel, U. et al. (1995) *Journal of Neurochemistry* 65, 2241-2250). *RAFTK* phosphorylation was inhibited following treatment by the PKC inhibitors Calphostin C or Staurosporine as well as BAPTA (an intracellular Ca<sup>2+</sup> chelator). Thus PKC is directly involved in *RAFTK* phosphorylation. BAPTA blocked PMA or SCF induced *RAFTK* phosphorylation, indicating that calcium was essential for PKC mediated *RAFTK* activation. Moreover, the calcium-ionophore (A23187) also stimulated *RAFTK* phosphorylation and was inhibited by PKC inhibitors Calphostin C and Staurosporine or by BAPTA, indicating again a role for PKC as a mediator of several signalling pathways including Ca<sup>2+</sup> in *RAFTK* phosphorylation.

The cytoskeleton is essential for many cellular functions including regulation of cell shape, flexibility, and adhesive properties (Hynes, R. O. (1992) *Cell* 69, 11-25; Juliano, R. L. and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585). Part of the cytoskeleton and plasma membrane form a region known as the focal adhesion (Lo, S. H. et al. (1994) *Bioessays* 16, 817-823). Focal adhesions are structures that form adherent contacts with the extracellular matrix. Proteins contained in the focal adhesion include talin, (-actinin, vinculin, paxillin, and other proteins (See, e.g., Tachibana, K. et al. (1995) *Journal of Experimental Medicine* 182, 1089-1099; Petch, L. A. et al. (1995) *Journal of Cell Science* 108, 1371-1379; Lewis, J. M. and Schwartz, M. A. (1995) *Molecular Biology of the Cell* 6, 151-160). The signal transduction pathways initiated by integrins involves cytoskeletal dependent activation of tyrosine kinases and phosphorylation of a number of substrates including

FAK protein (See, e.g., Juliano, R. L. and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585; Hamawy, M. M. et al. (1994) American Society of Microbiology, Washington DC, p. 235; Schaller, M. D. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196).

*RAFTK* is tyrosine phosphorylated upon fibronectin stimulation and co-localized with vinculin at "focal adhesion like structures" in CMK cells. *RAFTK* activation upon calcium ionophore (A23187) treatment or SCF or PMA stimulation of CMK cells is completely abolished in the presence of cytochalasin-D, which disrupts the cytoskeleton. These observations indicate that *RAFTK* is associated with the cytoskeleton and the integrity of the cytoskeleton is required for *RAFTK* phosphorylation, similar to that of FAK phosphorylation by integrins and external stimuli (Clark, E. A. and Brugge, J. S. (1995) *Science* 268, 233-239).

Constitutive *in vivo* association of *RAFTK* and paxillin was also observed. This effect of *RAFTK* on paxillin activation can promote paxillin binding to other SH2-domain containing proteins that might be involved in multiple signal transduction pathways.

Constitutive association between *RAFTK*, PKC- $\delta$  and paxillin *in vivo* was observed. This shows a role for *RAFTK* in linking and crosstalk between various signaling proteins localized in the cytosol and focal adhesion contacts. The RhoA-dependent assembly of focal adhesions in Swiss 3T3 cells was associated with increased tyrosine phosphorylation and the recruitment of both pp125FAK and PKC- $\delta$  to focal adhesions (Barry, S. T. and Critchley, D. R. (1994) *Journal of Cell Science* 107, 2033-2045). Association of PKC- $\delta$  with *RAFTK* and paxillin shows that phosphorylation of these components can be an important event in integrin mediated events in megakaryocytes.

The tyrosine kinase PYK2, which is identical to *RAFTK*, has been shown to be involved in calcium signalling and MAP kinases function in PC-12 neuronal cells (Lev, S. et al. (1995) *Nature* 376, 737). Stimulation of megakaryocytes with thrombin leads to tyrosine phosphorylation of *RAFTK* (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10). The evidence that *RAFTK* is involved in the c-kit growth factor signalling pathway in megakaryocytic cells as described in this study further supports the broad function of this kinase in a variety of signalling pathways.

**EXAMPLE 3: Characterization of *RAFTK*, a Novel Focal Adhesion Kinase, and Its Integrin-Dependent Phosphorylation and Activation in Megakaryocytes**

5

Biochemical characterization and functional analysis of the *RAFTK* protein was performed. Coexpression of *RAFTK* and FAK proteins in megakaryocytic cells and blood platelets was observed. Using a specific antibody to *RAFTK* and the monoclonal antibody 2A7 to FAK, FAK and *RAFTK* could be distinguished antigenically.

10

*RAFTK* had intrinsic tyrosine kinase and auto-kinase activities. It was phosphorylated on tyrosine in growing cultures of COS cells transfected with the pCDNAIII/flag-*RAFTK* expression vector containing the *RAFTK* cDNA ligated with the eight amino acid flag peptide sequence. Similar to FAK, dephosphorylation of *RAFTK* was observed when adherent transfected COS cells were detached. Phosphorylation was regained upon replating of these cells on the fibronectin-coated dishes. Analysis of tyrosine phosphorylated *RAFTK* from adherent transfected COS cells revealed that the Src homology 2 (SH2) domains of the Src and Fyn protein kinases as well as the Grb2 adaptor protein were able to specifically associate with *RAFTK*. Tyrosine phosphorylation of endogenous *RAFTK* was observed upon fibronectin induced activation of human megakaryocytic cells. Furthermore, colocalization of *RAFTK* protein with vinculin, a focal adhesion protein, was observed in "focal adhesion-like structures" in adherent CMK cells and in transfected pCDNAIII/flag-*RAFTK* COS cells upon fibronectin activation, by confocal microscopy.

15

20

25

These data show that *RAFTK* is a novel member of the FAK family, that it localizes to "focal adhesion-like structures" in CMK megakaryocytic cells, participates in integrin-mediated signaling pathways in megakaryocytes and is able to associate with the tyrosine kinases Src and Fyn as well as the adaptor protein Grb2 via SH2-phosphotyrosine interactions.

30

The following materials and methods were used to biochemically characterize and functionally analyze the *RAFTK* protein:

**35 Chemical and biological reagents**

Human fibronectin, Poly-L-Lysine (MW 70,000-150,000 Dalton) and geneticin (G418) were purchased from Sigma Chemical Co. (St. Louis, MO).

Monoclonal antibody to phosphotyrosine (PY20) was purchased from Zymed Laboratories, Inc. (South San Francisco, CA). Monoclonal antibody 2A7 and polyclonal antibody BC3 to pp125<sup>FAK</sup> were gifts from Dr. T. Parsons (University of Virginia, Charlottesville, VA). The 2A7 antibody was also purchased from Upstate Biotechnology Inc. (Lake Placid, NY). Polyclonal antibody 331 to FAK was a gift from Dr. S.K. Hanks (Vanderbilt University, Nashville, TN). Monoclonal antibody M5 to the flag peptide was purchased from Eastman Kodak Co. (New Haven, CT). rhGM-CSF was purchased from R & D Systems (Minneapolis, MN). Monoclonal anti-human antibody to vinculin was purchased from Sigma (St. Louis, MO).

### Cells and cell growth

COS cells were obtained from the American Type Tissue Culture (Rockville, MD). The cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Mediatech Co., Washington, D.C.) supplemented with 10% fetal calf serum (FCS) (Sigma), 2 mM glutamine, 100 U/ml of penicillin and 100 mg/ml of streptomycin. Megakaryocytic cell lines CMK, DAMI, CMS, Meg-01 and CMK11-5 were maintained in RPMI-1640 supplemented with 10% FCS, 2 mM glutamine, 100 U/ml of penicillin and 100 mg/ml of streptomycin as described previously (Sakaguchi M et al. (1991) *Blood* 77, 481; Komatsu N et al. (1989) *Blood* 74, 42; Greenberg S M et al. (1988) *Blood* 72, 1968; Avraham H et al. (1994) *Blood* 83, 2126).

### Expression of GST (Glutathione S-Transferase) fusion proteins

Oligonucleotides flanking various *RAFTK* domains and containing appropriate restriction sites were synthesized. The polymerase chain reaction (PCR) was used with *RAFTK* cDNA as a template to amplify the appropriate fragments. The DNA fragments encoding amino acid (a.a) residues 26-286 (N-*RAFTK*), 375-680 (KA-*RAFTK*), 375-1009 (KC-*RAFTK*) and 681-1009 (C-*RAFTK*) of *RAFTK* were amplified by the PCR technique and the sequences for these encoding regions were confirmed by DNA sequencing. The PCR products were precleaved with BamHI and EcoRI and were ligated into the pGEX-2T expression vector (Pharmacia Biotech, Inc., Piscataway, NJ) which had also been cleaved with BamHI and EcoRI. Competent *Escherichia coli* (E. coli) DH5a were transformed, and recombinant bacterial clones were screened by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of overexpressed fusion proteins and restriction enzyme analysis. GST-fusion proteins were produced by 1 mM isopropyl b-thiogalactopyranoside induction and purified by affinity chromatography on Glutathione-Sepharose beads (Pharmacia Biotech, Inc., Piscataway, NJ).

### Construction of pCDNAIII/flag and pCDNAIII/flag-*RAFTK*

The pCDNAIII/flag expression vector was constructed by inserting a short DNA fragment encoding a starting codon and an 8-a.a flag peptide into a pCDNAIII expression vector (Invitrogen Co., San Diego, CA) at HindIII and EcoRI sites. The two oligonucleotides used were: sense primer: 5'-AGC TTA TGG ACT ACA AGG ACG ACG ATG ACA GGG G-3' (SEQ ID NO:17); antisense primer: 5' AAT TCC CTT GTC ATC GTC GTC CTT ATG GTC CAT A-3' (SEQ ID NO:18). The cDNA encoding 1009 amino acids of human *RAFTK* was then subcloned in an EcoRI site located downstream of the flag sequences of the pCDNAIII/flag vector. The orientation and DNA sequences of the *RAFTK* cDNA were confirmed by DNA sequencing.

### Transfection of COS cells and analysis of *RAFTK* phosphorylation

COS cells were transfected by the calcium phosphate method using pCDNAIII/flag-*RAFTK* or pCDNAIII/flag expression vectors according to the manufacturer's protocol (Invitrogen Co., San Diego, CA). The transfected cells were starved in serum-free DMEM for 4-6 hr, harvested by phosphate-buffered saline (PBS) containing 2 mM EDTA and washed with PBS twice. The cells ( $1.5 \times 10^6$  per 60 mm dish) were then plated onto fibronectin (5.0  $\mu$ g/ml) or Poly-L-Lysine (5.2  $\mu$ g/ml) coated dishes at 37°C for various times (20 or 40 min). Adherent cells were lysed in 1 ml of RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 10  $\mu$ g/ml of aprotinin, leupeptin and pepstatin, and 1 mM  $\text{Na}_3\text{VO}_4$ ). Analysis of *RAFTK* phosphorylation was performed as described further.

### Immunoprecipitation

Total cell lysates (TCL) were clarified by centrifugation at  $10,000 \times g$  for 10 min. Protein concentrations were determined by the protein assay (Bio-Rad Laboratories, Hercules, CA) and were standardized to equal concentrations of protein prior to immunoprecipitation. Identical amounts of protein from each sample were precleared by incubation with protein G-Sepharose CL-4B (Sigma Co., St. Louis, MO) for 1 hr at 4°C. After the removal of protein G-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as described below for each experiment for 4 hr or overnight at 4°C. Immunoprecipitation of the antigen-antibody complex was accomplished by incubation for 2 hr at 4°C with 25  $\mu$ l of protein G-Sepharose. Non-specific bound

proteins were removed by washing the Sepharose beads three times with HNTG buffer (20 mM Hepes, pH 7.4, 150 mM NaCl, 10% glycerol and 0.1% Triton X-100). Bound proteins were solubilized in 20  $\mu$ l of 2 x Laemmli buffer and further analyzed by immunoblotting.

5

#### **Endogenous *RAFTK* phosphorylation upon integrin activation**

20 x 10<sup>6</sup> CMK cells were starved in serum-free RPMI-1640 culture medium overnight. The cells were divided into 4 parts: one portion was replated onto fibronectin-coated (8  $\mu$ g/cm<sup>2</sup>) dishes; one portion onto collagen-coated (8  $\mu$ g/cm<sup>2</sup>) dishes; one portion onto Poly-L-Lysine (5  $\mu$ g/cm<sup>2</sup>) dishes; and one portion was kept in suspension. After 1 hr replating, the medium was aspirated and adherent cells were gently and quickly washed with ice-cold PBS. The cells were lysed in 1 ml RIPA buffer and cleared by centrifugation for 10 min at 10,000 rpm. 800  $\mu$ g of TCL was incubated overnight at 40C with 10  $\mu$ l of R-4250, followed by immunoprecipitation with protein-A-Sepharose beads for 2 hr. The complexes were washed with HNTG buffer 3 times and then analyzed by Western blot analysis.

15

#### **Immunoblot**

A defined amount of protein lysate was combined with the same volume of Laemmli loading buffer and boiled for 2 min. In the case of immunoprecipitates, 20  $\mu$ l of 2 x Laemmli loading buffer was added. Samples were separated and analyzed by 8% SDS-PAGE gel and then transferred to nitrocellulose membranes. The membranes were blocked with 5% bovine serum albumin (BSA) in PBS containing 0.1% Tween-20 (Boehringer Mannheim Co., Indianapolis, IN) and probed with primary antibody for 1 hr at room temperature according to the enhanced chemiluminescent (ECL) protocol (Amersham Corp., Arlington Heights, IL). Immunoreactive bands were visualized using HRP-conjugated secondary antibodies and ECL reagents (Amersham Corp., Arlington Heights, IL).

25

#### ***In vitro* kinase assay**

Immunoprecipitated complexes were washed twice with RIPA buffer and once in kinase buffer (20 mM Hepes; pH 7.4, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, and 100  $\mu$ M Na<sub>3</sub>VO<sub>4</sub>). The washed immune complexes or a defined amount of purified GST-fusion proteins were suspended in 20  $\mu$ l of kinase buffer and then [g-<sup>32</sup>P]ATP was added up to 250  $\mu$ Ci/ml in the presence of 25  $\mu$ g of poly (Glu:Tyr) (4:1, 20 to 50 Kd) (Sigma Chemical Co., St. Louis, MO), at RT for 15 min. The reaction

35

was stopped by the addition of 10 mM Hepes (pH 7.4)/10 mM EDTA. The labeled substrates were analyzed by SDS-PAGE and autoradiography.

### ***In vitro* SH2 association assays**

- 5           *In vitro* association experiments were carried out with GST-fusion proteins containing the SH2 domains of Fyn and Grb2 (a generous gift of Dr. L. Cantley, Harvard Medical School, Boston, MA) and Src (a generous gift of Dr. T. Pawson, Mount Sinai Research Institute, Canada). TCL (1 mg) was incubated for 120 min at 4°C with 10 µg of GST-fusion proteins coupled to Glutathione-Sepharose beads in the  
10       absence or presence of the indicated amount of synthetic peptide. The beads were washed three times with Tris-buffer saline, and proteins were separated by 8% SDS-PAGE.

### **Antibodies**

- 15           Anti-*RAFTK* antiserum was obtained from New Zealand white rabbits immunized with a bacterially expressed GST-fusion protein containing the C-terminal (681-1009 amino acid residues) of human *RAFTK* cDNA. The sera were titrated by ELISA against the GST-*RAFTK* C-terminus fusion protein. The serum R-4250 exhibited the highest titer (1:256,000) and was found to react specifically with the  
20       GST-fusion proteins containing the C-terminus of *RAFTK*. This serum was used in subsequent experiments.

### **Confocal microscopy**

- (1) *Cell staining*: Cultured CMK megakaryocytes or transfected COS cells  
25       were plated overnight on glass coverslips coated with fibronectin (5-8 µg/cm<sup>2</sup>). Adherent cells were fixed with 2% paraformaldehyde in PBS, pH 7.4 for at least 15 min, permeabilized in PBS containing 0.2% Triton X-100 for 2 min, washed in PBS, and placed in a blocking buffer containing PBS, 3% BSA and 1% normal goat serum for 1 hr. Adherent cells were incubated with anti-vinculin (IgG1 mouse anti-human,  
30       1:200 dilution) and *RAFTK* (rabbit anti-*RAFTK*, 1:100 dilution) antibodies for 1 hr. Cells were washed in PBS, incubated with FITC conjugated (anti-mouse IgG), and Texas red conjugated (anti-rabbit IgG) antibodies for 1 hr.

- (2) *Microscopy*: Stained cells were briefly washed in PBS, then sealed in  
microwell chambers containing Vectashield antifade and examined using a Sarastro  
35       2000 confocal laser scanning microscope (CLSM) (Molecular Dynamics, Sunnyvale, CA) fitted with a 25 mW argon-ion laser. The microscope was configured for dual channel fluorescent imaging with: 488/514 nm excitation, 535 nm primary



beamsplitter, 10% laser transmission, 18 mW laser power. A 595 nm secondary beamsplitter passed fluorescent light emitted from *RAFTK*-labeled (Texas red)-cells to a photomultiplier tube fitted with a 600 long pass filter. Short wavelength light (<595 nm) emitted from the vinculin-labeled focal adhesion plaques (FITC) was directed to a second photomultiplier tube fitted with a  $540 \pm 15$  nm band pass filter and simultaneously recorded. Image pairs were subjected to a 2-D median filter to reduce background noise, then examined as color composite images with *RAFTK* appearing red, and vinculin green. Focal adhesion plaques expressing both *RAFTK* and vinculin appeared yellow-orange.

(3) *Image analysis*: In some cases, the distribution of *RAFTK* and vinculin were determined using a quantitative image analysis procedure in which 2-D pixel intensity histograms from both the *RAFTK* and vinculin images were compared using ImageSpace (Molecular Dynamics, Sunnyvale, CA) software. Image analysis was performed on pairs of images to determine the area occupied by *RAFTK* and vinculin within the CMK cells. Pixel dimensions of all micrographs were  $0.17 \mu\text{m}$ , and pixel intensities ranged from 0-255 intensity units.

The pixel intensity range corresponding to the cell cytoplasm was determined separately for each image. Cell background was found to be within a range of 4-73 pixel intensity units. The noncellular background composed of surrounding media and slide surface was determined to be 0-3 pixel intensity units. Cells containing fluorescent stained *RAFTK* were observed to have pixel intensities from 74-255, whereas pixel intensities for vinculin ranged between 86-254 units. Pixel intensities corresponding to either *RAFTK* or vinculin were applied to a 2-D histogram of pixel intensities with *RAFTK* (X-axis), plotted against vinculin (Y-axis).

Pixels unique to each either *RAFTK* or vinculin were identified on the 2-D histogram and converted into a binary section. This section was applied as a mask over the original image to produce area measurements on a per cell basis. This procedure was repeatedly used to measure the area occupied by both *RAFTK* and vinculin for both cell types.

**The *RAFTK* gene is highly conserved during species evolution and is coexpressed with FAK in several hematopoietic cells.**

*RAFTK* shares about 65% similarity of its amino acid sequence with that of FAK, suggesting that *RAFTK* and FAK may have some common features. In order to determine whether *RAFTK* and FAK may be coexpressed in the same cells, total RNA from different megakaryocytic cell lines was prepared and Northern blot analysis was performed using a human *RAFTK*-specific probe and a human FAK-specific probe.

Northern blot analysis showed that all tested megakaryocytic cell lines coexpressed both *RAFTK*- and FAK- specific transcripts of 4.5 Kb.

Coexistence of *RAFTK* and FAK proteins was detected by Western blot analysis in the CMK megakaryocytic cell line and in blood platelets using the polyclonal antibody 4250 to the C-terminal *RAFTK* and the monoclonal antibody 2A7 to FAK. Total lysates from  $5 \times 10^8$  platelets of  $10^6$  CMK cells were prepared and protein concentrations were determined. An equal amount of protein was immunoprecipitated by polyclonal antiserum 4250 for *RAFTK* or normal rabbit serum (NRS) or immunoprecipitated with the monoclonal antibody 2A7 to FAK or control irrelevant monoclonal antibody with the same isotype. The immunocomplexes were resolved by SDS-PAGE and then immunoblotted with antibody 4250 (1:1000) or immunoblotted with 2A7 (1:500). Antibodies 4250 and 2A7 were shown to be specific to *RAFTK* and FAK, respectively, as described below. Taken together, these results demonstrated that *RAFTK* and FAK were coexpressed in these hematopoietic cells.

***RAFTK* is antigenically distinct from, but related to FAK.**

In order to investigate whether *RAFTK* and FAK were antigenically related, Three DNA fragments encoding N-terminal (N-*RAFTK*), kinase catalytic (KA-*RAFTK*) and C-terminal (C-*RAFTK*) domains of *RAFTK* were subcloned into the pGEX-2T expression vector. In addition, the full length cDNA was ligated with a sequence encoding an eight amino acid flag peptide and subcloned into a pCDNAIII vector.

Expression of the GST-fusion proteins, N-*RAFTK*, KA-*RAFTK* and C-*RAFTK* as well as GST protein was analyzed by Western blot using antibodies to GST and FAK. Only the C-terminal GST-*RAFTK* (amino acids 681-1009) was recognized by polyclonal antibodies BC3 and 331, respectively, which were generated against the C-terminal of the FAK protein (Schaller M D et al. (1992) *Proc Natl Acad Sci USA* 89, 5192; Hanks S K et al. (1992) *Proc Natl Acad Sci USA* 89, 8487). This result showed that *RAFTK* is antigenically related to FAK. However, the monoclonal antibody 2A7 against FAK did not crossreact with any of the three fusion proteins, suggesting that *RAFTK* may be antigenically different from FAK.

To further confirm these findings, a polyclonal antiserum 4250 by rabbit immunization with the C-terminal *RAFTK* GST-fusion protein was generated. This antiserum, like the monoclonal antibody M5 to the flag peptide and polyclonal Ab BC3, specifically recognized the M5-immunoprecipitated flag-*RAFTK* protein expressed in transfected COS cells. To verify 2A7 antibody specificity, antibody 2A7

but not 4250 or M5, was shown to recognize a 2A7-immunoprecipitated 125 Kd protein in either vector alone or flag-*RAFTK* transfected COS cells; this protein appears to represent the endogenous FAK. Similarly, *RAFTK* immunoprecipitated by antiserum 4250 from CMK megakaryocytic cell lysates was not recognized by the  
5 monoclonal antibody 2A7 to FAK and FAK immunoprecipitated by the monoclonal antibody 2A7 from CMK cell lysates was not recognized by *RAFTK*-specific antibodies. Taken together, these data show that FAK and *RAFTK* are distinguishable antigenically while being related members of the FAK family.

10 ***RAFTK* has intrinsic protein tyrosine kinase and auto-kinase activities.**

Enzymatic activity of a protein tyrosine kinase is essential for its role in signal transduction. To assess intrinsic tyrosine kinase activity, kinase activity of the purified N-*RAFTK*, KA-*RAFTK*, KC-*RAFTK*, and C-*RAFTK* GST-fusion proteins in *in vitro* assays in which poly (Glu:Tyr) (4:1) was used as an exogenous substrate was tested.  
15 The results showed that the KC-*RAFTK* fusion protein possessed kinase activity. Similar data were obtained by testing total recombinant bacterial cell lysates, as only the KC-*RAFTK* construct had kinase activity. These results demonstrate that *RAFTK* has intrinsic PTK activity.

Autophosphorylation of FAK is an initial step in cell response to stimuli and  
20 provides a dock for the association of Src-family kinases (Weng Z et al. (1993) *J Biol Chem* 268, 14956; Schaller M D et al. (1994) *Mol Cell Biol* 14, 1680). In order to test *RAFTK* auto-kinase activity, the incorporation of <sup>32</sup>P into the purified GST-fusion proteins containing either the kinase catalytic domain, C-terminal domain or kinase plus C-terminal domain was tested in the absence of exogenous substrate. In  
25 agreement with the results from the *in vitro* kinase activity assay, only the kinase catalytic domain plus C-terminal domain of *RAFTK* appeared capable of autophosphorylating itself (MW 100 Kd). The additional low molecular weight proteins might represent the proteolytic isoforms of the KC-*RAFTK* GST-fusion protein which were found during recombinant *E. coli* growth and induction, as  
30 detected by antibodies to the GST protein. These results show that *RAFTK* possesses auto-kinase activity.

***RAFTK* is involved in integrin-mediated signaling in transiently transfected COS cells and CMK megakaryocytic cells.**

35 Since *RAFTK* is structurally similar to FAK which plays a central role in integrin-mediated signaling pathways, *RAFTK* phosphorylation upon integrin engagement was studied. Tyrosine phosphorylation and kinase activity of flag-

*RAFTK* in transiently transfected COS cells were analyzed. When transfected COS cells were grown on plastic culture dishes, flag-*RAFTK* protein was phosphorylated. Detachment of the transfected cells by 1 mM EDTA in PBS resulted in a significant decrease in the level of *RAFTK* phosphorylation; replating the cells onto fibronectin-coated dishes increased the phosphorylation of flag-*RAFTK* in a time dependent manner. In contrast, replating the cells onto Poly-L-Lysine coated dishes had no effect on the phosphorylation of flag-*RAFTK*. These data show that the phosphorylation of flag-*RAFTK* is modulated by integrin interaction with fibronectin.

To assess whether the kinase activity of *RAFTK* was stimulated by integrin activation, the flag-*RAFTK* from detached cells, cells attached to fibronectin or cells attached to Poly-L-Lysine were analyzed. While flag-*RAFTK* from fibronectin-coated dishes was markedly increased in its kinase activity, no increased kinase activity was found in flag-*RAFTK* from Poly-L-Lysine coated dishes. The flag-*RAFTK* in the detached cells retained a very low level of activity.

In order to elucidate whether or not endogenous *RAFTK* is responsive to integrin activation (like FAK), a detailed analysis of integrin-mediated signaling of endogenous *RAFTK* in megakaryocytes was also performed. After starvation, CMK cells were lysed in RIPA buffer as a control, or replated onto fibronectin, collagen or Poly-L-Lysine coated dishes for 1 hr. The adherent cells were quickly washed and lysed in RIPA buffer. A total of 1.2 mg TCL for each sample was immunoprecipitated with anti-*RAFTK* serum (R-4250). After washing, the immunocomplexes were divided into three parts: equivalent of 1 mg TCL for phosphorylation analysis; equivalent of 180 ug TCL for autophosphorylation assay; and equivalent of 20 ug TCL for kinase assay. The phosphorylation of *RAFTK* was significantly increased in CMK megakaryocytic cells adherent to fibronectin or collagen coated dishes, while no phosphorylation of *RAFTK* was observed in nonadherent CMK cells or CMK cells grown onto Poly-L-Lysine dishes. These data show that *RAFTK* can be activated by integrin engagement in CMK cells.

To assess whether the auto-kinase activity of *RAFTK* was stimulated by integrin activation, endogenous *RAFTK* in CMK cells untreated or stimulated with collagen, fibronectin, or Poly-L-Lysine was analyzed. There was an increase in autophosphorylation of *RAFTK* upon the collagen or fibronectin adhesion of CMK cells. In contrast, a very low level of autophosphorylation in CMK cells adherent to Poly-L-Lysine and no autophosphorylation activity in untreated CMK cells was observed.

In order to evaluate the intrinsic kinase activity of endogenous *RAFTK* in CMK cells, the kinase activity of endogenous *RAFTK* in CMK cells adherent to

fibronectin, collagen, or Poly-L-Lysine was tested. Strong phosphorylation of the exogenous substrate poly (Glu:Tyr) (4:1) by *RAFTK* was observed when CMK cells were adherent to collagen or fibronectin, while low kinase activity of *RAFTK* was detected when CMK cells were adherent to Poly-L-Lysine or in suspension. These results, together with the *in vitro* kinase experiments using the GST-*RAFTK* fusion proteins, demonstrated that *RAFTK* increased its intrinsic PTK and auto-kinase activities upon integrin activation.

**Association of phosphorylated *RAFTK* with Src-family kinases and Grb2 adaptor protein via SH2-phosphotyrosine interactions.**

The tyrosine phosphorylation of several signaling proteins provides a specific dock for the association of SH2-containing proteins. To test whether phosphorylated *RAFTK* associated with the Src-family kinases and the Grb2 protein, GST-SH2 fusion proteins from Fyn, Src, Grb2, N-terminal and C-terminal p85 subunits of PI3-K (NP85 and CP85), respectively were purified to homogeneity. These proteins were incubated with TCL from flag-*RAFTK* transfected COS cells adherent on fibronectin-coated dishes. Associated complexes were analyzed by Western blot analysis. GST-SH2 fusion proteins from Fyn, Src and Grb2 appeared to be specifically associated with phosphorylated *RAFTK* while GST protein alone, GST-NP85SH2 and GST-CP85SH2 did not bind to phosphorylated *RAFTK*. In order to test whether such association was enhanced specifically by fibronectin treatment, *RAFTK* protein from transfected COS cells in suspension was compared to that of transfected COS cells adherent on fibronectin in its association with SrcSH2, FynSH2 and Grb2SH2 GST-fusion proteins. Cells adherent on fibronectin significantly increased the association of *RAFTK* with SrcSH2, FynSH2 and Grb2SH2 GST proteins. These results demonstrate that *RAFTK* strengthens its ability to bind Src, Fyn and Grb2 molecules after cellular integrin activation.

Since Grb2SH2 GST-fusion proteins seemed to be associated less with *RAFTK* than the SrcSH2 and FynSH2 GST-fusion proteins, this specificity was further confirmed by abolishing this association with increasing concentrations of a phosphotyrosine synthetic peptide (VpYLNVMEL) corresponding to amino acids 880-887 of *RAFTK*. Taken together, these results demonstrated that tyrosine-phosphorylated *RAFTK* has the ability to specifically associate with Src and Fyn kinases and Grb2 protein in a SH2-dependent manner.

**Localization of *RAFTK* to "focal adhesion-like structures" of CMK cells and transfected COS cells.**

To analyze whether endogenous *RAFTK* in CMK cells was localized to focal  
5 adhesion structures or to cell-cell contacts, a detailed analysis using confocal  
microscopy was performed. In addition, localization of *RAFTK* in transfected COS  
cells was determined. Purified R-4250 antibodies which specifically detect *RAFTK*  
were used. Confocal image analysis of immunostained CMK cells and transfected  
10 COS cells adherent to the fibronectin substrate reveal "focal adhesion-like structures"  
adjacent to the glass coverslip surface. Confocal image analysis showed greater than  
90% of the *RAFTK* was colocalized with vinculin under these conditions. Cells  
immunostained with *RAFTK* revealed punctate areas of staining near the basal  
surfaces of cells which were adherent to fibronectin for 12 hrs. Cells adherent to  
15 fibronectin for 1 hr revealed similar colocalization to basal cell surfaces, however,  
focal adhesion plaques were not well developed at this time point.

In this study, the biochemical characterization and functional analysis of a  
novel signaling molecule, *RAFTK*, which is abundantly expressed in megakaryocytes,  
platelets and brain tissue (Avraham S et al. (1995) *J Biol Chem* 270, 27742) is  
described. The results show that *RAFTK*, like FAK, possesses intrinsic protein  
20 tyrosine kinase and auto-kinase activities; is coexpressed with FAK in  
megakaryocytic cells and platelets; and is immunologically related to, but distinct  
from FAK. The phosphorylation and kinase activity of *RAFTK* were stimulated by  
integrin engagement. Phosphorylated *RAFTK* was able to specifically bind to Src-  
family kinases and the Grb2 adaptor protein via an apparent phosphotyrosine-SH2  
25 interaction. These data demonstrate that *RAFTK* is a novel member of the FAK  
family and shares structural, immunological, enzymatic, and functional features with  
FAK.

Fibronectin stimulation increased *RAFTK* tyrosine phosphorylation when an  
epitope tagged *RAFTK* was expressed into COS cells. In addition, the  
30 phosphorylation of *RAFTK* was significantly increased in CMK cells adherent onto  
fibronectin or collagen coated dishes, while no phosphorylation of *RAFTK* was  
observed in untreated CMK cells or CMK cells grown onto Poly-L-Lysine coated  
dishes. These results clearly show that the phosphorylation of endogenous *RAFTK* is  
modulated by integrin interaction with fibronectin or collagen in CMK cells. In  
35 addition, microscopic imaging of CMK cells and transfected pCDNAIII/flag-*RAFTK*  
COS cells following double-staining with vinculin and *RAFTK* revealed  
colocalization of the *RAFTK* protein with vinculin in "focal adhesion-like structures"

in CMK and transfected COS cells treated with fibronectin. It is important to note that all published studies involving focal adhesion sites were done in adherent cells (such as 3T3 cells) where the staining of focal adhesion structures is in a conventional punctate pattern. Megakaryocytic as well as CMK cells are cells grown in suspension and lack the typical focal adhesion structures. Therefore, these confocal studies were done in transfected COS cells as well as CMK cells grown in fibronectin and the stained structures are called "focal adhesion-like structures". Furthermore, the colocalization of *RAFTK* with vinculin, a well-known focal adhesion protein, was performed in CMK and transfected COS cells. The colocalization of both proteins is about 90%. The microscopic studies support the biochemical evidence that *RAFTK* is localized to "focal adhesion-like structures" surrounded by vinculin in adherent CMK cells and transfected COS cells.

These studies on *RAFTK* localization were performed in transfected COS cells (with a flag-*RAFTK* construct) and in CMK cells. The plane of focus was set to be within the focal adhesion plaque region at 0.2 microns above the coverslip surface to exclude cell-cell contact. Furthermore, the tyrosine phosphorylation of endogenous *RAFTK* was observed upon collagen stimulation of platelets. However, in PC12 cells, *RAFTK* was not phosphorylated upon collagen treatment. These observations suggest that *RAFTK* phosphorylation upon integrin-mediated signaling is dependent on cell types and integrin forms, indicating a cell type specific signaling event.

Amino acid and DNA sequence homology studies showed that *RAFTK* is most closely related to FAK, sharing 65% similarity (Avraham S et al. (1995) *J Biol Chem* 270, 27742). Such high similarity between the proteins suggested that *RAFTK* and FAK may have similar molecular structural conformations. This prediction was supported by antigenic crossreactivity studies in which two polyclonal antibodies to FAK recognized the C-terminal GST-fusion protein of *RAFTK* as well as flag-*RAFTK*. *RAFTK* appears to be structurally distinct from FAK. A further comparison of *RAFTK* and FAK in different regions indicated that NH2 and COOH terminal domains have more divergence than the kinase domain. Such divergence may account for the failure of monoclonal antibody 2A7 to FAK and polyclonal antibody 4250 to *RAFTK* to recognize common epitopes.

Since FAK-family kinases lack SH2 and SH3 domains, the regulation of tyrosine phosphorylation plays a critical role in protein-protein interactions during signal transduction. More than half (20/35) of the tyrosine residues of *RAFTK* are highly conserved in the FAK molecule (Avraham S et al. (1995) *J Biol Chem* 270, 27742). Importantly, two of these residues in FAK were identified as being phosphorylated and sequentially bound to the SH2 domains of the Src-family kinases

and the Grb2 adaptor protein: Tyr<sup>397</sup> (Schaller M D et al. (1994) *Mol Cell Biol* 14, 1680) and Tyr<sup>925</sup> (Schlaepfer D.D. et al. (1994) *Nature* 372, 786), respectively. The sequences downstream of these two phosphotyrosines are consistent with the prediction that the Src SH2 domain preferentially binds to the phosphotyrosine sequence pYAEI, whereas Grb2 binds to pYENV (Songyang Z et al. (1993) *Cell* 72, 767). Equivalent phosphotyrosine sequences were found at Tyr<sup>402</sup> and Tyr<sup>881</sup> of *RAFTK*, however, glutamic acid next to Tyr<sup>881</sup> was substituted with leucine. These results demonstrated that the SH2 domains of Src, Fyn and Grb2 were able to specifically associate with tyrosine phosphorylated *RAFTK* from fibronectin-activated COS cells. The leucine substituted next to phosphotyrosine-881 did not change its specificity of binding for Grb2. However, unlike FAK, *RAFTK* contains more than one potential binding site for Src-family tyrosine kinases. These results show that *RAFTK*, like FAK, is a substrate for Src kinases which are required for FAK family kinase mediated signaling pathway(s).

Tyrosine kinase and auto-kinase activities are essential for FAK to initiate its downstream signaling pathway. Because *RAFTK* has very large NH<sub>2</sub> and COOH domains, it is important to preclude the possibility that any other *RAFTK*-associated kinase(s) may contribute to the enzymatic activity in the *in vitro* kinase assay. Therefore, purified GST-*RAFTK* fusion proteins produced from recombinant bacteria were tested for kinase activity. The data described herein demonstrated that *RAFTK*, like FAK, possesses intrinsic kinase and auto-kinase activities. However, it is interesting to observe that such activities may require not only the kinase catalytic domain but also the COOH domain, which would differ from that observed with FAK produced in mammalian cells (Eide B L et al. (1995) *Mol Cell Biol* 15, 2819; Chan P Y et al. (1994) *J Biol Chem* 269, 20567). It is unclear whether the different observations imply different characteristics of FAK and *RAFTK* or are due to different expression systems.

**EXAMPLE 4:           Activation of a Novel Related Focal Adhesion Tyrosine Kinase (*RAFTK*) During an Early Phase of Platelet Activation by an Integrin GpIIb-IIIa Independent Mechanism**

***RAFTK* is rapidly tyrosine phosphorylated in thrombin-stimulated platelets,**

To elucidate the role of *RAFTK* activation in platelets the effect of thrombin on *RAFTK* tyrosine phosphorylation was studied. Thrombin induced a dose and time dependent phosphorylation of *RAFTK* in platelets. Treatment of platelets with 0.05,



0.1 or 0.25 U/ml of thrombin induced a basal level of *RAFTK* phosphorylation similar to the resting or unstimulated platelets. An increase, but equal levels of *RAFTK* phosphorylation was induced with 0.5 or 1 U/ml of thrombin, while 2 U/ml of thrombin induced highest levels of *RAFTK* phosphorylation.

- 5           A time course of thrombin stimulation in platelets showed a rapid induction of *RAFTK* phosphorylation. The resting or unstimulated platelets showed very low basal level of phosphorylation. Activation was observed as early as 10 seconds post-thrombin stimulation, reaching a maximum at 2 minutes and tapering off by 10 minutes. These results showed that thrombin stimulation of platelets induces *RAFTK* phosphorylation in a time and concentration dependent manner.
- 10

***RAFTK* is an endogenous substrate for Calpain.**

- During the course of platelet activation, agonist induced activation of calpain (Fox, J. E. B. et al. (1990) *Blood* 76, 2510-2519; Fox, J. E. et al. (1991) *J. Biol. Chem.* 266, 13289-13295; Saido, T. et al. (1993) *J. Biol. Chem.* 268, 7422-7426) and limited proteolysis of some specific substrates (Ando, Y. et al. (1987) *Biochem. Biophys. Res. Commun.* 144, 484-490; Tsujinaka, T. et al. (1982) *Thromb. Res.* 28, 149-156; Oda, A. et al. (1993) *J. Biol. Chem.* 268, 12603-12608; Frangioni, J. V. et al. (1993) *EMBO J.* 12, 4843-4856) has been reported. Calpain constitutes most of the calcium dependent protease activity in platelets (Fox, J. E. B. et al. (1990) *Blood* 76, 2510-2519; Oda, A. et al. (1993) *J. Biol. Chem.* 268, 12603-12608). Thrombin is one of the agonists that causes activation of calpain in platelets.
- 15
- 20

- Despite the use of equal number of platelets ( $5 \times 10^8$  /ml) for immunoprecipitations, a dose and time dependent decrease in *RAFTK* protein levels after thrombin stimulation was consistently observed. The decrease in the protein levels could be due to protein degradation or protein processing. Since *RAFTK* undergoes processing upon thrombin stimulation, an examination was performed to determine if calpain is involved in this process by using a specific membrane permeable calpain inhibitor, calpeptin (Thujinaka, T. et al. (1988) *Biochem. Biophys. Res. Commun.* 153, 1201-1208). When platelets were activated by thrombin, no differences were seen in the levels of *RAFTK* phosphorylation between calpeptin pretreated or thrombin alone treated platelets. There was no detectable phosphorylation in untreated or resting platelets. However, when *RAFTK* protein levels were examined on the same immunoblot, thrombin treated sample showed cleavage of *RAFTK* where as calpeptin treated platelets showed a complete blockage in *RAFTK* degradation. The level of the *RAFTK* protein in calpeptin treated platelets was equal to the level in the resting or unstimulated platelets. Furthermore, *RAFTK*
- 25
- 30
- 35

processing was observed in a time dependent manner in response to the pharmacological activator of calpain, A23187, and to the physiological activator of calpain collagen. The characteristics of inhibition of degradation of *RAFTK* were consistent with the involvement of calcium dependent neutral cysteine protease, since  
5 degradation of *RAFTK* occurred at neutral pH and was inhibited by cysteine protease inhibitor, calpeptin. These results provide evidence that agonist induced *RAFTK* processing is mediated through the activation of calpain.

10 ***RAFTK* phosphorylation is independent of aggregation and occurs early, during platelet shape change.**

The earliest platelet response induced by physiological agonists involves a change in shape from flat discs into compact spheres followed by secretion of granular contents. The later phase of platelet response is aggregation when large platelet aggregates are formed. Since *RAFTK* phosphorylation is induced early (10 seconds)  
15 after thrombin stimulation, it was investigated if aggregation is a prerequisite for *RAFTK* phosphorylation. In order to prevent aggregation, platelets were activated by thrombin in the absence of stirring. A time course of thrombin induced platelets showed that *RAFTK* is rapidly phosphorylated as early as 10 seconds in the presence  
20 or absence of stirring (aggregation). However, the signal declined faster upon stirring, where as it persisted longer in the absence of stirring (10 min). The levels of *RAFTK* protein remained equal in the absence of stirring and the levels decreased only in the presence of stirring, indicating that *RAFTK* cleavage is dependent on aggregation, while *RAFTK* phosphorylation does not require aggregation. However some  
25 aggregation can occur in the absence of stirring because of the close proximity of the platelets ( $5 \times 10^8$ /ml). Therefore, in order to further confirm that *RAFTK* phosphorylation does not require aggregation, platelets were pretreated with an aggregation inhibitor, RDGS tetrapeptide, before activation by an agonist. Treatment with RGDS peptide allows shape change but prevents aggregation by blocking the  
30 interaction of fibrinogen with the integrin GpIIb/IIIa. Addition of thrombin to platelets pretreated with increasing concentrations of RGDS or RGES (mock) peptides showed no change in the levels of tyrosine phosphorylation of *RAFTK*. Thus, these results confirm that *RAFTK* phosphorylation does not require aggregation and it occurs during platelet shape change. Furthermore, consistent with previous findings, proteolytic processing of *RAFTK* was inhibited by RGDS treatment (absence of  
35 aggregation), but it was not inhibited when platelets were treated by RGES (presence of aggregation) or with thrombin alone. The phosphorylation of *RAFTK* correlated well with an early wave of tyrosine phosphorylation especially of proteins pp60<sup>src</sup>

and pp72<sup>syk</sup> (Clark, E. A., and Brugge, J. S. (1993) *Mol. Cell. Biol.* 13, 1863-1871; Clark, E.A. et al. (1994) *J. Biol. Chem.* 269, 288859-28864), but it precedes the wave of tyrosine phosphorylation of its family member, pp125<sup>FAK</sup> known to be dependent on platelet aggregation (Clark, E. A. et al. (1994) *J. Biol. Chem.* 269, 288859-28864).

**Activation of *RAFTK* is not dependent on the integrin, GpIIb/IIIa.**

GpIIb/IIIa is a major integrin receptor which plays an important role in adhesive events critical in clot formation by binding to fibrinogen and von Willebrand factor matrix proteins (Fox, J. E. B. et al. (1993) *J. Biol. Chem.* 268, 25973-25984). Activation of FAK was found to be mediated through the integrin GpIIb/IIIa (Lipfert, L. et al. (1992) *J. Cell Biol.* 119, 905-912). Since *RAFTK* is a member of FAK subfamily, whether *RAFTK* and FAK have similar mechanisms of regulation was investigated. Tyrosine phosphorylation of *RAFTK* was studied under conditions that specifically induce or inhibit fibrinogen binding to this receptor. The monoclonal antibody 7E3 binds to GpIIb/IIIa and blocks fibrinogen binding (Coller, B. S. et al. (1989) *Circulation* 80, 1766-1774). Incubation of platelets with 7E3 for 20 minutes prior to stirring, followed by addition of thrombin did not inhibit tyrosine phosphorylation of *RAFTK* while tyrosine phosphorylation of FAK was inhibited under the same conditions. Pretreatment with a (control) monoclonal antibody 6D1, specific for collagen receptor GpIa/IIb did not alter thrombin induced phosphorylation of *RAFTK* or FAK. Preincubation of platelets with 7E3, 6D1 or buffer alone without thrombin stimulation showed no phosphorylation of *RAFTK* or FAK. These results showed that phosphorylation of *RAFTK* was not dependent either on fibrinogen binding to GpIIb/IIIa or platelet aggregation, and therefore, the phosphorylation of *RAFTK* in platelets is not solely regulated through ligand occupancy of the integrin GpIIb/IIIa. Furthermore it is interesting to note that preincubation with 7E3 plus thrombin stimulation, but not 6D1 prevented proteolytic processing of *RAFTK*. These results indicate that *RAFTK* phosphorylation is not dependent on GpIIb/IIIa, but the proteolytic processing of *RAFTK* is dependent GpIIb/IIIa. To further confirm that phosphorylation of *RAFTK* is not dependent on the activation (crosslinking) of GpIIb/IIIa, fibrinogen binding to the stirred platelets was initiated by an anti-B3 antibody Fab fragment (anti-LIBS6), in the absence of an agonist. This antibody renders GpIIb/IIIa competent to bind fibrinogen, but it does not itself cause detectable platelet activation (Huang, M.-M. et al. (1993) *J. Cell Biol.* 122, 473-483). *RAFTK* was not phosphorylated when platelets were stirred or unstirred with anti-LIBS6 and fibrinogen, in platelets treated with fibrinogen alone or in resting platelets despite

containing the protein. *RAFTK* was however activated when platelets were treated with thrombin (positive control). Thus these studies indicate that activation of *RAFTK* does not require crosslinking of GpIIb/IIIa receptors on the platelet surface.

5     ***RAFTK* activation is regulated by actin polymerization.**

Thrombin stimulation in platelets leads to actin polymerization and causes dramatic rearrangements of the cytoskeleton thereby inducing the formation of focal-contact like areas (Furman, M. I. et al. (1993) *Thromb. Haemostasis* 70, 229-232). It was also examined whether phosphorylation of *RAFTK* was affected by agents that  
10     disrupt the actin cytoskeleton. Platelets pretreated with cytochalasin D block agonist induced actin polymerization but do not inhibit platelet aggregation. Pretreatment with cytochalasin D inhibited tyrosine phosphorylation of *RAFTK* in thrombin stimulated platelets. However the level of inhibition was not 100%, suggesting that the actin-dependent cytoskeletal interactions effected partially phosphorylation of  
15     *RAFTK*. Furthermore proteolytic processing of *RAFTK* is not inhibited in cytochalasin D treated platelets.

***RAFTK* is activated by multiple platelet agonists.**

*RAFTK* has been found to be activated by thrombin, calcium ionophore,  
20     collagen, and the combination of ADP plus epinephrine.

25     **EXAMPLE 5:           Identification of a Novel Signal Transduction Pathway in Human Macrophages Mediated by the Related Adhesion Focal Tyrosine Kinase (*RAFTK*)**

*RAFTK*, a novel non-receptor protein kinase, has been shown to be involved in focal adhesion signal transduction pathways in neuronal PC12, megakaryocytes and platelets. Because focal adhesions may modulate cytoskeleton function and thereby  
30     alter phagocytosis, cell migration, and adhesion in macrophages, the role of *RAFTK* signaling in these cells was investigated. *RAFTK* was abundantly expressed in THP1 monocytic cells as well as primary alveolar and peripheral blood derived macrophages. Phorbol diester stimulation of THP1 cells increased tyrosine phosphorylation of *RAFTK* by 2.5 minutes. Similar increases in phosphorylation were  
35     detected within 1 minute after CSF-1/MCSF stimulation. While *RAFTK* was phosphorylated with similar kinetics in peripheral blood derived macrophages, alveolar macrophages showed high constitutive phosphorylation levels which

decreased over increased time after treatment with either PMA or CSF-1/M-CSF. Immunoprecipitation analysis identified constitutive associations between *RAFTK* and the cytoskeleton protein paxillin and the signaling molecule PI-3 kinase. However, both these molecules appear disassociate from *RAFTK* at the peak time of phosphorylation after PMA or CSF-1/M-CSF stimulation. *RAFTK* was also found to preferentially associate with the amino terminus-SH3 domain of the Grb2 adaptor protein in THP1 cells. Furthermore the CSF-1/M-CSF receptor *fms* and *RAFTK* appear to associate in response to CSF-1/M-CSF treatment of THP1 cells. These data demonstrate that *RAFTK* participates in macrophage signal transduction pathways mediated by CSF-1/M-CSF.

With this background, it was investigated whether *RAFTK* was expressed in human monocyte-macrophages and whether it participated in CSF-1/M-CSF induced signaling. In parallel, the effects of treatment of monocyte-macrophages with the known chemical activator phorbol diester PMA were tested. It was observed that *RAFTK* was robustly expressed in both peripheral blood derived monocyte-macrophages as well as tissue derived alveolar macrophages. Moreover, it was activated upon treatment of mononuclear phagocytes with CSF-1/M-CSF or PMA and associated with other signaling molecules and the cytoskeletal protein paxillin. These observations provide new data on CSF-1/M-CSF signaling and molecules which may contribute to focal adhesion formation in cells of this lineage.

The following materials and methods were used to identify the novel signal transduction in human macrophages mediated by *RAFTK*:

## 25 **Cells and cell cultures**

The permanent human monocyte-macrophage cell line THP-1 was obtained from the American Type Culture Collection (ATCC) and shown to be mycoplasma free prior to expansion in culture. The cells were carried in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal calf serum (FCS), 2mM glutamine, sodium pyruvate, non-essential amino acids and 50µg/ml penicillin and 50µg/ml streptomycin. Primary human peripheral blood monocyte-macrophages (MMs) were obtained by phlebotomy of normal volunteers after obtaining informed consent and isolated by ficoll hypaque density centrifugation as previously described (Boyum, A. (1968) *Scand. J. Lab Invest.* 21, (Suppl. 97). Mms plated on plastic dishes for 24 hours were shaken at 200 RPM for 15 minutes and washed 3X with HBSS to remove non-adherent cells. The adherent population of cells used for subsequent studies as previously described (Kharbanda, S. et al. (1995) *Proc. Nat.*

*Acad. Sci. USA*, 92, 6132-6136). Alveolar macrophages (AMs) were obtained by bronchoalveolar lavage of normal non-smoking volunteers after informed consent was obtained and using standard procedures.

## 5 Reagents and materials

Phorbol 12-myristate 13-acetate (PMA) was obtained from Sigma (St. Louis, MO) and dissolved in dimethyl sulfoxide and stored at -20°C until use. Recombinant human CSF-1/M-CSF was kindly provided by Genetics Institute.

The monoclonal antibodies against phosphotyrosine (4G10), PI-3 kinase p85 regulatory subunit, Grb2, and paxillin and the polyclonal rabbit antisera antibody to the human c-fms receptor were obtained from Upstate Biotechnology. Specific polyclonal antibodies to *RAFTK* were generated by immunizing New Zealand White rabbits with a bacterially expressed fusion protein consisting of GST and the carboxy terminus (amino acids 681-1009) of human *RAFTK* cDNA subcloned into the pGEX-2T expression vector as described (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751). High titer *RAFTK* antiserum (R-4250)) was employed in subsequent experiments and was shown to be specific and not cross reactive with FAK in prior experiments (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751).

The Grb2 and PI-3 kinase GST fusion proteins were obtained from Santa Cruz Biotechnology. Electrophoresis reagents and nitrocellulose membranes were obtained from Bio-Rad Laboratories (Hercules, CA). All other chemicals including the protease inhibitors pepstatin, antipain, chymostatin, leupeptin, aprotinin, and alpha 1 antitrypsin were obtained from Sigma (St. Louis, MO). Because endotoxin is known to alter monocyte-macrophage function, all media and reagents were shown to be free of endotoxin contamination by Limulus endotoxin assay (Sigma Chemical) prior to using in cell cultures.

## Signal transduction studies

Cells were initially starved in DMEM with 0.5% FCS and stimulated in HBSS at density of  $5 \times 10^6$ /ml for varying time periods at 37°C with PMA (2nM-200nM/ml) or CSF-1/M-CSF (10U-10,000U/ml). After stimulation  $20 \times 10^6$  cells were microfuged for 10 seconds and lysed in 1 ml of ice cold modified RIPA buffer (50mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150mM NaCl, 1mM PMSF, 10 ug/ml of pepstatin, antipain, chymostatin, leupeptin, aprotinin, alpha 1 antitrypsin, 10mM sodium fluoride and 10 mM sodium pyrophosphate). Total cell lysates (TCL) were clarified by centrifugation at  $10,000 \times g$  for 10 min. Protein concentrations were determined by protein assay (BioRad Laboratories).

### Immunoprecipitation and Western blot analysis

For immunoprecipitation studies, identical amounts of protein from each sample were clarified by incubation with protein A-Sepharose CL-4B (Pharmacia Biotech) for 1 h at 4°C. Following the removal of protein A-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as detailed below for each experiment for 4 h or overnight at 4°C. Immunoprecipitations of the antibody-antigen complexes were performed by incubation for 2 h at 4°C with 75 µl of protein A-Sepharose (10% suspension). Non-specific bound proteins were removed by washing the Sepharose beads three times with modified RIPA buffer and one time with Phosphate buffered saline (PBS). Bound proteins were solubilized in 30 µl of 2 X Laemmli buffer and further analyzed by immunoblotting. Samples were separated on 7.5% SDS-PAGE and then transferred to nitrocellulose membranes. The membranes were blocked with 5% non-fat milk protein and probed with primary antibody for 3 h at RT or 4°C overnight. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) system (Amersham Corp., Arlington Heights, IL). Blots were stripped (2% SDS, 62.5mM Tris, 100mM Beta Mercaptoethanol) for 30 minutes at 50°C and washed in TBS-T for 60 minutes before blocking and re-probing with primary antibodies.

### GST-fusion protein binding studies

GST-fusion protein Grb2-SH3 N-terminal domain, Grb2-SH3 C-terminal domain, Grb2-SH2 domain, and PI-3 Kinase-SH3 domain of the p85 regulatory subunit were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). For the binding experiments, 1 mg of cell lysate was mixed with 5 g of GST-fusion protein and incubated for 1 h at 4°C on a rotatory shaker. 50 µl of glutathione Sepharose 4B beads (Pharmacia Biotech) were added to preabsorb the complex. Following incubation for 3 h at 4°C on a rotatory shaker, the beads were centrifuged and washed three times with modified RIPA buffer. The bound proteins were eluted by boiling in Laemmli sample buffer and subjected to SDS-PAGE on 7.5% gel and Western Blot analysis.

### *RAFTK* is expressed and phosphorylated in human monocyte-macrophages.

To further characterize signaling pathways in human MMs involved in their growth, differentiation and function, the permanent monocyte-macrophage cell line THP-1 as well as primary peripheral blood derived MMs or AMs were used as a model. Analysis by immunoprecipitation revealed abundant *RAFTK* protein in these

cells. There appeared to be low levels of constitutive phosphorylation of *RAFTK* in these cells under unstimulated culture conditions.

Then it was addressed whether various stimuli associated with mononuclear phagocyte activation modulated *RAFTK* phosphorylation. Preliminary experiments determined that 1000U/ml was optimal for CSF-1/M-CSF and 20nM/ml was optimal for PMA stimulation of *RAFTK* in THP1 cells and primary macrophage cultures. An increase in the tyrosine phosphorylation of *RAFTK* was specifically observed in THP1 cells following PMA and CSF-1/M-CSF treatment. The membrane was then stripped and reprobed with anti-*RAFTK* antibody to confirm that equivalent amounts of *RAFTK* were loaded in each lane.

To determine the time course of tyrosine phosphorylation of *RAFTK*, THP1 cells were stimulated with PMA or CSF-1/M-CSF. Phosphotyrosine levels in *RAFTK* immunoprecipitates peaked at 2.5 minutes and declined by 5 minutes. However, phosphotyrosine levels again increase by 7.5 minutes and decline after 10 minutes. There were not any changes in *RAFTK* protein levels to explain these fluctuations in phosphotyrosine levels. The membrane was then stripped and reprobed with anti-*RAFTK* antibody to confirm that equivalent amounts of *RAFTK* were loaded in each lane.

CSF-1/M-CSF treatment of THP1 cells resulted in maximum tyrosine phosphorylation of *RAFTK* within 0.5 minutes to 1 minute which declined by 2.5 minutes. Similar to the case following PMA stimulation, the phosphotyrosine levels on *RAFTK* also appeared to increase by 10 minutes. Longer stimulation times confirmed the fluctuation of *RAFTK* tyrosine phosphorylation.

CSF-1/M-CSF stimulation of MM resulted in peak tyrosine phosphorylation on *RAFTK* by 1 minute which gradually decreased over time. *RAFTK* in AM appeared to have a high constitutive degree of phosphorylation which, in response to CSF-1/M-CSF stimulation, increased slightly by 1 minute and gradually decreased over time. Anti-*RAFTK* immunoblotting of *RAFTK* immunoprecipitates showed the ~120Kd phosphoprotein corresponded to the *RAFTK* protein. Depending on the resolution of the gels, *RAFTK* was seen to migrate either as a single band or as a doublet.

*RAFTK* associates with the signaling molecules PI-3 kinase and Grb-2. Because *RAFTK*, like FAK, acts as a platform kinase site for the coalescence of signaling and adaptor molecules at sites of focal adhesions, *RAFTK* immunoblots were examined for associating co-precipitating proteins. A specific association of *RAFTK* with PI-3 kinase, an important enzyme in modulating of phosphoinositol signaling (Auger, K. R., and Cantley, L. C. (1991) *Cancer Cells* 3, 263-270) was



observed. This association was confirmed by incubating THP1 cell lysates using a PI-3 kinase GST-fusion protein, immunoprecipitated with glutathione-conjugated beads and detecting the bound proteins by anti-*RAFTK* immunoblotting. Time course studies using either CSF-1/M-CSF or PMA treatment demonstrated that the PI-3 kinase/*RAFTK* association fluctuates. While the PI-3 kinase signal appears to weaken after one minute its association with *RAFTK* strengthens after 5 minutes of stimulation. A similar pattern is detected with longer stimulation times.

Because *RAFTK* has been shown to associate with various SH2 and SH3 domain-containing proteins, the ability of *RAFTK* to form *in vitro* complexes with the adaptor molecule Grb2 was then examined. Grb2 is generally capable of associating with signaling molecules either through one of its two SH3 or through its SH2 domain. Studies were thus performed to determine which of the Grb2 regions may mediate its interaction with *RAFTK*. THP1 cell lysates incubated with GST fusion proteins corresponding to the amino terminus SH3, the SH2 and the carboxyl terminus SH3 were immunoprecipitated with glutathione-conjugated beads and the bound proteins were detected by anti-*RAFTK* immunoblotting. *RAFTK* strongly associates with the Grb2 amino terminus-SH3 domain. There was some intermediate interactions between the SH2 domain and *RAFTK* molecule in CSF-1/M-CSF stimulated THP1 cells.

#### ***RAFTK* associates with the c-fms receptor upon mononuclear phagocyte cell activation with CSF-1/M-CSF.**

Because CSF-1/M-CSF stimulation of THP1 cells or primary macrophages appeared to have very rapid effects on *RAFTK* phosphorylation, whether *RAFTK* may directly associate with the c-fms receptor was examined. A specific association of *RAFTK* with the c-fms receptor upon CSF-1/M-CSF treatment of cells was observed. Associations were detected in reciprocal blotting experiments of THP1 cell lysates either immunoprecipitated with *RAFTK* followed by c-fms immunoblot or c-fms immunoprecipitation followed by *RAFTK* immunoblot. No association was detected between these molecules in unstimulated or PMA stimulated THP1 cells.

#### ***RAFTK* associates with the cytoskeletal protein paxillin in mononuclear phagocyte.**

Following the observation that *RAFTK* may be coimmunoprecipitated with molecules previously characterized as components of MM signaling pathways, it was desirable to determine whether certain cytoskeletal molecules in such cells also associate with this novel kinase. Using specific antibodies to *RAFTK* or paxillin, an

important molecule involved in focal adhesions, a clear association of these two molecules in THP1 cells was found. Similar to PI-3 kinase, the strength of paxillin's constitutive association with *RAFTK* transiently fluctuated after 1 minute stimulation by either CSF-1/M-CSF or PMA stimulation. The paxillin *RAFTK* association  
5 appeared to return to constitutive levels after 5 minutes stimulation.

These studies indicate that human mononuclear phagocytes, including peripheral blood derived MMs and tissue derived AMs, express *RAFTK*, a recently identified signaling molecule that is a member of the FAK family. *RAFTK* appeared to participate in certain previously described signaling pathways following activation  
10 of these cells. Treatment with CSF-1/M-CSF and phosphorylation of the cognate c-fms receptor revealed robust phosphorylation of *RAFTK* in both the model THP-1 cell line as well as in primary macrophages. Parallel studies using the chemical activator PMA also revealed phosphorylation of *RAFTK* in macrophages in a time and concentration dependent manner.

The phosphorylation *RAFTK* was found to result in association with several well characterized components of cell signal pathways, including the enzyme PI-3 kinase and the adaptor molecule Grb2. Grb2 is an adaptor protein that has the capacity to link with a number of kinases and substrates and functions to facilitate signaling through creation of physical associations of such partners in enzymatic  
20 reactions (Pawson, T. (1995) *Nature* 373, 573-580). PI-3 kinase appears to modulate phosphoinositol metabolism in a variety of cell types, including mononuclear phagocytes, and appears to be an important component of tyrosine kinase-regulated signaling pathways that lead to cell proliferation (Gold, M. R. et al. (1994) *J. Biol. Chem.* 269, 5403-5412). CSF-1/M-CSF has been reported to induce the direct  
25 association of the p85a subunit of PI-3 kinase with the SH2 domain of Grb2 and Grb2-Sos complexes, supporting its role upstream of the Ras signaling pathway in monocytes (Saleem, A., et al. (1995) *J. Biol. Chem.* 270, 10380-10388). In addition, PI-3 kinase activation and the production of its metabolites has been suggested to be an upstream activator of calcium-independent form of PKC (Herrera-Velit, P. and  
30 Reiner, N. E. (1996) *J. Immunology* 156, 1157-1165).

These observations on *RAFTK* show that this recently identified signaling molecule plays a variety of roles in transduction of MM signaling. It provides a missing link in prior studies of CSF-1/M-CSF induced integrin expression and the subsequent formation of focal adhesion contacts, as reported by De Nichilio and  
35 Yamada (De Nichilo, M. D. and Yamada, K. M. (1996) *J. Biol. Chem.* 271, 11016-11022).

Taken together, the data presented herein contributes to the model of CSF-1/M-CSF mediated signaling in mononuclear phagocytes. c-Fms has been reported to form associations with Shc, Grb2 and Sos1 in myeloid cells suggesting c-Fms signals through the Ras pathway (Liobin, M. N. et al. (1994) *Molecular and Cellular Biology* 14, 5682-5691). The finding that *RAFTK* associated with c-Fms in CSF-1/M-CSF stimulated THP1 cells shows that c-Fms also signals through focal adhesion contacts in concert with integrin binding. Thus, macrophages, like megakaryocytes, prominently utilize *RAFTK* in cytokine mediated pathways of activation that are linked to focal contact formation. The confluence of *RAFTK* and other kinases and cytoskeletal molecules provides a platform for the interactions of signaling molecules and adaptor proteins that regulate cell morphology to finely control certain components of the immune response such as adhesion or migration.

**EXAMPLE 6:           *RAFTK*, a Novel Member of the Focal Adhesion Kinase Family, Participates in T-Cell Receptor Signal Transduction**

It has been found that *RAFTK* is constitutively expressed in human T-cells and is rapidly phosphorylated upon the activation of the T-cell receptor (TCR) for antigen. This activation results in an increase in the association of *RAFTK* with the Src cytoplasmic tyrosine kinase Fyn. *RAFTK* also associates with the SH2 domain of Grb2 and with the cytoskeletal protein paxillin. The tyrosine phosphorylation of *RAFTK* following T-cell receptor-mediated stimulation was reduced by the pre-treatment of cells with cytochalasin D, indicating the role of the cytoskeleton in this process. These observations show that *RAFTK* participates in T-cell receptor signaling and acts to link signals from the cell surface to the cytoskeleton and thereby effect the host immune response.

It was observed that *RAFTK* is phosphorylated in response to the activation of certain integrins in megakaryocytes (Li, J. et al. (1996) *Blood* 88, 417-428) and B-lymphocytes. The induced phosphorylation of *RAFTK* via calcium-mediated ion channel pathways was shown first in PC-12 pheochromocytoma cells (Lev, S. et al. (1995) *Nature* 376, 737-745) and subsequently in megakaryocytes.

Thus, *RAFTK* is expressed in human T-lymphocytes and participates in signaling events triggered by the ligation of the TCR/CD3 complex. Several of the interacting molecules that associate with *RAFTK* in human T-cells, including the cytoskeletal protein paxillin have been characterized.

The following materials and methods were used to determine the participation of *RAFTK* in T-cell receptor signal transduction:

### Cells and cell cultures

5           The permanent human T-cell lines Jurkat and H9 were obtained from the American Type Culture Collection (ATCC) and shown to be mycoplasma-free prior to their expansion in culture. The cells were carried in Dulbecco's Modified Eagle's Medium (DMEM) with 10% fetal calf serum (FCS), 2 mM glutamine, 50 g/mL penicillin and 50 g/ml streptomycin. Primary human peripheral blood lymphocytes  
10 (PBLs) were obtained by phlebotomy of normal volunteers after obtaining their informed consent and isolated by ficoll hypaque density centrifugation as previously described (Boyum, A. (1968) *Scand. J. Lab Invest.* 21, (Suppl. 97)). The anti-CD3 producing hybridoma (OKT-3) was obtained from ATCC and grown in Iscove's modified Dulbecco's medium with 20% FCS. For antibody production, cells were  
15 grown in serum-free and protein-free hybridoma medium (Sigma, St. Louis, MO) containing Nutridoma-HU 1% (Boehringer Mannheim, Indianapolis, IN).

### Reagents and materials

20           The lectin phytohemagglutinin (PHA) was obtained from Pharmacia Biotech (Piscataway, NJ). The nitrocellulose membrane was obtained from Bio-Rad Laboratories (Hercules, CA). The anti-phosphotyrosine monoclonal antibody (4G10, IgG2a) was obtained from Upstate Biotechnology Inc. (Lake Placid, NY). The anti-CD3 antibody X35 was obtained from Immunotech (Marseille, France), and OKT-3 was purified from OKT-3 producing hybridoma supernatants on protein A-Sepharose  
25 columns. Antibodies to Fyn were obtained from Santa Cruz Biotechnology (Santa Cruz, CA) and Transduction Laboratories (Lexington, KY). Anti-paxillin antibody was obtained from ICN Biomedical Inc. (Costa Mesa, CA). Specific polyclonal antibodies to *RAFTK* were generated by immunizing New Zealand White rabbits with a bacterially expressed fusion protein consisting of GST and the carboxy terminus  
30 (amino acids 681-1009) of the human *RAFTK* cDNA subcloned into the pGEX-2T expression vector as described (Li, J. et al. (1996) *Blood* 88, 417-428). The sera were titrated against the GST-*RAFTK* C-terminus fusion protein by ELISA and the serum (R-4250) which revealed the highest titer (1:256,000) was employed in the subsequent experiments. This antiserum was shown to be specific and not crossreactive with  
35 FAK in prior experiments (Li, J. et al. (1996) *Blood* 88, 417-428). Electrophoresis

reagents were obtained from Bio-Rad Laboratories (Hercules, CA). The protease inhibitors leupeptin, aprotinin, and alpha 1 antitrypsin and all other reagents were obtained from Sigma Co. (St. Louis, MO).

## 5 Stimulation of cells

Cells were washed twice with Hanks Balanced Salt Solution, Gibco (Grand Island, NY) and resuspended at  $5 \times 10^6$  cells/ml in DMEM medium. Cells were stimulated with either PHA (10 g/ml) or CD3 antibodies X35 (10 g/ml) or OKT-3 (10 g/ml) at 37°C for various time periods. In some experiments, cells were pretreated with cytochalasin D (2  $\mu$ M) for 60 min at 37°C before stimulation. After stimulation,  $20 \times 10^6$  cells were microfuged for 10 seconds and lysed in 1 ml of modified RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 10 g/ml of aprotinin, leupeptin and pepstatin, 10 mM sodium vanadate, 10 mM sodium fluoride and 10 mM sodium pyrophosphate). Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by protein assay (Bio-Rad Laboratories).

## Immunoprecipitation and Western blot analysis

For immunoprecipitation studies, identical amounts of protein from each sample were clarified by incubation with protein A-Sepharose CL-4B (Pharmacia Biotech) for 1 h at 4°C. Following the removal of protein A-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as detailed below for each experiment for 4 h or overnight at 4°C. Immunoprecipitations of the antibody-antigen complexes were performed by incubation for 2 h at 4°C with 75  $\mu$ l of protein A-Sepharose (10% suspension). Non-specific bound proteins were removed by washing the Sepharose beads three times with modified RIPA buffer and one time with Phosphate buffered saline (PBS). Bound proteins were solubilized in 40  $\mu$ l of 2 X Laemmli buffer and further analyzed by immunoblotting. Samples were separated on 7.5% SDS-PAGE and then transferred to nitrocellulose membranes. The membranes were blocked with 5% non-fat milk protein and probed with primary antibody for 3 h at room temperature or 4°C overnight. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) system (Amersham Corp., Arlington Heights, IL).

## 35 GST-fusion protein binding studies

GST-fusion protein Grb2-SH3 N-terminal domain, Grb2-SH3 C-terminal domain, Grb2-SH2 domain, and Fyn-SH2 and -SH3 domains were purchased from

Santa Cruz Biotechnology (Santa Cruz, CA). For the binding experiments, 1 mg of cell lysate was mixed with 5 g of GST-fusion protein and incubated for 1 h at 4°C on a rotatory shaker. 50 µl of glutathione Sepharose 4B beads (Pharmacia Biotech) were added to preabsorb the complex. Following incubation for 3 h at 4°C on a rotatory shaker, the beads were centrifuged and washed three times with modified RIPA buffer. The bound proteins were eluted by boiling in Laemmli sample buffer and subjected to SDS-PAGE on 7.5% gel and Western Blot analysis.

***RAFTK* is expressed in human T-lymphocytes and is phosphorylated upon T-cell activation.**

To further characterize signaling pathways in human T-cells involved in the immune response, Two permanent T-cell lines, Jurkat and H9, were utilized as well as primary circulating PBLs. Analysis by immunoblot or immunoprecipitation revealed abundant *RAFTK* protein in these cells.

The stimulation of human T-cell lines with T-cell receptor ligation induces the tyrosine phosphorylation of a phosphoprotein around 115 Kd (Motto, D. G. et al. (1994) *Journal of Biological Chemistry* 269, 21608-21613; Hsi, E. D. et al. (1988) *The Journal of Biological Chemistry* 264, 10836-10842). It was investigated whether various stimuli associated with such T-cell activation modulated the phosphorylation of *RAFTK*, which has a deduced molecular weight of ~120 Kd. An increase in the tyrosine phosphorylation of *RAFTK* could be specifically observed in the T-cell lines Jurkat or H9 following T-cell receptor ligation or treatment with the lectin PHA. The membrane was then stripped and reprobed with anti-*RAFTK* antibody to confirm that equivalent amounts of *RAFTK* were loaded in each lane. Stimulation of primary circulating PBLs with anti-T-cell receptor antibody also induced an increase in the tyrosine phosphorylation of *RAFTK*.

To determine the time course of tyrosine phosphorylation of *RAFTK*, Jurkat cells were stimulated with anti-T-cell receptor antibody X35 or OKT-3 or with the lectin PHA. Ligation of the TCR/CD3 by monoclonal antibody X35 or OKT-3 reached a maximum by 2.5-5 min, and declined thereafter. PHA stimulation resulted in an increased tyrosine phosphorylation by 5 min which declined slightly thereafter with substantial phosphorylation still detectable at 20 min. Anti-*RAFTK* immunoblotting of anti-*RAFTK* immunoprecipitates showed that the ~115 Kd phosphotyrosine polypeptide corresponds to the *RAFTK* protein. Depending on the resolution of the gels, *RAFTK* was seen to migrate either as a single band or as a doublet.

***RAFTK* associates with the signaling molecules Fyn and Grb-2.**

To further characterize the role that *RAFTK* plays in T-cell signaling following activation via TCR/CD3 ligation, as well as the other stimulatory pathways activated by the lectin PHA, coimmunoprecipitation studies followed by immunoblotting were performed. A specific association of *RAFTK* with Fyn, a src family kinase which is known to be capable of associating with TCR was observed. A small fraction of Fyn was readily detected as associating with *RAFTK* prior to the TCR/CD3 activation of Jurkat cells and this association increased following stimulation.

The ability of *RAFTK* to form *in vitro* complexes with various SH2 and SH3 domain-containing proteins was then examined. For this purpose, GST-fusion proteins were added to the lysates of the stimulated Jurkat cells, the complexes were immunoprecipitated with glutathione-conjugated beads and the bound proteins were detected by anti-*RAFTK* immunoblotting. Sepharose beads containing the GST-Grb2-SH2 domain and the GST-Fyn-SH2 domain bound *RAFTK* from the activated T-cell lysates. Beads containing only GST, GST-N-terminal Grb2-SH3, GST-C-terminal Grb2-SH3 or GST-Fyn-SH3 failed to bind *RAFTK*, indicating the specificity of these interactions. These data clearly show that the stable interaction between *RAFTK* and the SH2 domain of Fyn can be mimicked *in vitro* and provide additional evidence that the SH2 domain may be the principal determinant of *RAFTK* binding to Fyn *in vivo*.

***RAFTK* associates with the cytoskeletal protein paxillin.**

Following the observation that *RAFTK* may be coimmunoprecipitated with molecules previously characterized as components of the TCR signaling pathways, it was desirable to determine whether certain cytoskeletal molecules in T-cells may also associate with this novel kinase. Using specific antibodies to *RAFTK* or paxillin, a constitutive association of these two molecules was found.

To further investigate the cytoskeletal dependence of the tyrosine phosphorylation of *RAFTK*, Jurkat cells, prior to TCR stimulation, were pre-incubated for 60 min at 37°C with media alone or with cytochalasin D. The phosphorylation of *RAFTK* was reduced following the cytochalasin D treatment of cells.

These studies show that *RAFTK*, a novel signaling molecule that appears to be a member of the FAK family, is present in human T-lymphocytes and participates in signaling pathways following T-cell activation. Following the ligation of the TCR/CD3 there was a robust phosphorylation of *RAFTK* in both the model permanent T-cell lines, Jurkat and H9, as well as in primary PBLs. Parallel studies using other T-cell activators, specifically the lectin PHA, revealed a similar activation of *RAFTK* in a time and concentration dependent manner.

It is noteworthy that following activation, *RAFTK* was found to be associated with several well-characterized components of TCR/CD3 signaling pathways, including Fyn and Grb2. Fyn is known to be capable of associating with the TCR/CD3 complex, and is believed to play an important role in initiating the changes in phosphorylation that lead to further downstream signaling. This role has been most clearly demonstrated in studies showing the impaired development of CD4+CD8+ thymocytes from double mutant mice rendered null for Fyn and FAK through homologous recombination (Kanazawa, S. et al. (1996) *Blood* 87, 865-870). Also, transgenic thymocytes from mice overexpressing Fyn were hyperstimulatable, and overexpression of a catalytically inactive form of Fyn substantially inhibited TCR-mediated activation in otherwise normal thymocytes (Cooke, M. P. et al. (1991) *Cell* 65, 281-291). Grb2 is a well-characterized adaptor molecule that seems capable of associating with a number of kinases and substrates and may also act to facilitate signaling through the enhancement of the physical association of such partners in enzymatic reactions (See, e.g., Li, N. et al. (1993) *Nature* 363, 85-88; Koch, C. A. et al. (1991) *Science* 252, 668-674; Pawson, T., and Gish, G. D. (1992) *Cell* 71, 359-362). Shc and Grb2 have also been shown to play important roles in T-cell signaling (Motto, D. G. et al. (1994) *Journal of Biological Chemistry* 269, 21608-21613; Meisner, H. et al. (1995) *Molecular & Cellular Biology* 15, 3571-3578; Fukazawa, T. et al. (1995) *Journal of Biological Chemistry* 270, 19141-19150).

These observations on *RAFTK* show that this novel signaling molecule can play a variety of roles in the transduction of T-cell signaling. The confluence of signaling molecules and cytoskeletal components provides a platform for the regulated interactions of kinases and substrates and lead to important changes in cell morphology that enable other aspects of the immune response such as adhesion or migration. Analogy with work in adherent mesenchymal cells shows that the formation of the so-called focal adhesions facilitate the creation of these platforms and mediate cell attachment and transduction of signals (See, e.g., Richardson, A. and Parsons, J. T. (1995) *Bioessays* 17, 229; Schaller, M. D. et al. (1992) *Proceedings of the National Academy of Sciences of the United States of America* 89, 5192-5196; Clark, E. A. and Brugge, J. S. (1995) *Science* 268, 233-239). Relatively little is known about similar mechanisms in hematopoietic cells like T-lymphocytes. Recently, another member of the FAK family, termed FAK B, was identified. Initial studies indicated that FAK B may associate with ZAP-70, an intracytoplasmic protein tyrosine kinase also capable of associating with TCR (Kanner, S. B. et al. (1994) *Proceedings of the National Academy of Sciences of the United States of America* 91, 10484-10487).



There is relatively limited information available on the convergence of protein tyrosine kinases and cytoskeletal elements in T-lymphocytes. Several T-cell surface structures, including CD11a/CD18 and CD44, associate with the cytoskeleton upon receptor cross-linking. Recently, the interaction of the chain of TCR with the actin cytoskeleton upon T-cell activation was demonstrated (Valitutti, S. et al. (1995) *Journal of Experimental Medicine* 181, 577-584). These results revealed that *RAFTK* co-associates with paxillin, a major component of the cytoskeleton. Furthermore, the pre-treatment of cells with cytochalasin D results in the reduced tyrosine phosphorylation of *RAFTK* upon T-cell receptor activation. This response shows that *RAFTK* phosphorylation requires the formation of a cytoskeletal complex which provides a foundation for the interactions and compartmentalization of kinases and substrates.

**EXAMPLE 7: Cytokine Signalling Through the Novel Tyrosine Kinase**  
***RAFTK* in Kaposi's Sarcoma Cells**

Considerable data indicate a role of endogenous and exogenous cytokines in the pathogenesis of Kaposi's sarcoma (KS). A number of growth factors including basic FGF, VEGF, oncostatin M (OSM), IL-6, and TNF- $\alpha$  have been reported to promote KS cell growth. A novel tyrosine kinase receptor, FLT-4, was found to be present on normal lymphatic endothelium and robustly expressed in KS cells. Moreover, the recently identified ligand VRP for the FLT-4 receptor results in signalling in KS cells. Signal transduction pathways following receptor engagement by these diverse cytokines that belong to different receptor families was studied. KS cells expressed a recently identified focal adhesion kinase termed *RAFTK* which is believed to coordinate surface signals from cytokine and integrin receptors with the cytoskeleton. *RAFTK* was phosphorylated in KS cells following treatment with b-FGF, OSM, IL-6, VEGF, VRP, or TNF- $\alpha$ . Following *RAFTK* activation by these cytokines, there was enhanced association of *RAFTK* with the cytoskeletal protein paxillin. This association appeared to be mediated through the C-terminal domain of *RAFTK* based on studies using GST-fusion proteins of different *RAFTK* domains. A novel surface receptor FLT-4 expressed on KS cells, as well as a novel intracytoplasmic tyrosine kinase *RAFTK* have been identified. Treatment with diverse cytokines previously reported to potentiate KS cell growth all led to phosphorylation of *RAFTK* and its association with the cytoskeletal protein paxillin. These

observations suggest that inhibition of *RAFTK* may allow for disruption of a common pathway important in KS cell growth and could be clinically exploited as an anti-neoplastic strategy.

5 Kaposi's sarcoma (KS) is the most frequent neoplasm arising among patients with the acquired immune deficiency syndrome (AIDS). The cell of origin of the tumor is believed to be from lymphatic endothelium (Dictor, M. (1988) *Lymphology*. 21, 53-60; Dorfman, R. F. (1988) *Lymphology*. 21, 45-52). Etiological factors implicated in KS include the recently discovered human herpes virus 8 (HHV-8)/Kaposi's sarcoma herpes virus (KSHV) and TAT, the soluble transcriptional  
10 activator of HIV (See, e.g., Chang, Y. et al. (1994) *Science*. 266, 1865-1869; Zhong, W. et al. (1996) *Proc. Natl. Acad. Sci. USA*. 93, 6641-6646; Huang, Y. Q. et al. (1996) *J. Clin. Invest.* 97, 2803-2806). Considerable data indicate a role of endogenous and exogenous cytokines in the pathogenesis of KS (See, e.g., Ba, G. et al. (1992) *J. Immunol.* 149, 3727-3734; Buonaguro, L. et al. (1992) *J. Virol.* 66, 7159-  
15 7167; Cai, J., et al. (1994) *Am. J. Pathol.* 145, 74-79). Growth factors such as basic fibroblast growth factor (b-FGF) and vascular endothelial growth factor (VEGF) which are known to stimulate mitogenesis of certain types of endothelium, as well as oncostatin M, interleukin-6 (IL-6), and tumor necrosis factor alpha (TNF- $\alpha$ ), which are elaborated during inflammatory conditions, have been implicated in promoting KS  
20 cell growth.

Defining the signal transduction pathways which may be utilized by cytokines which appear to stimulate KS growth provides an opportunity for rational and targeted therapeutic intervention against this neoplasm. One issue that is immediately  
apparent is that the cytokines described to date as promoting KS belong to distinctly  
25 different families as defined by their receptors. VEGF and b-FGF receptors are of the protein tyrosine kinase family, oncostatin M and IL-6 utilize a common gp130 subunit, and TNF- $\alpha$  receptors are of the Fas/apoptosis CD95 family. Cognizant of this diversity, the signalling pathways triggered by cytokine treatment in permanent human KS cells *in vitro* have been characterized and a common molecule sought  
30 among the diverse pathways. KS cells express *RAFTK*.

In KS cells, treatment with cytokines of different families, including b-FGF, oncostatin M, IL-6, VEGF, and TNF- $\alpha$ , all led to phosphorylation of *RAFTK*. After cytokine treatment, *RAFTK* is found to associate with the cytoskeletal protein paxillin. This observation has been extended and focused on the tyrosine kinase receptor  
35 termed FLT-4, which has been found in fetal and adult lymphatic endothelium (Kaipainen A. et al. (1995) *Proc. Natl. Acad. Sci. USA*. 92, 3566-3570; Kaipainen A. et al. (1993) *J. Exp. Med.* 178, 2077-2088; Pajusola K. et al. (1993) *Oncogene*. 8,

2931-2937). KS cells express the FLT-4 receptor, and treatment with its newly discovered ligand called VEGF related protein (VRP or VEGF-C) (Lee, J. et al. (1996) *Proc. Natl. Acad. Sci. USA*. 93, 1988-1992; Joukov V., et al. (1996) *EMBO J*. 15, 290-298) again results in *RAFTK* phosphorylation.

5

The following materials and methods were used to study cytokine signalling through *RAFTK* in Kaposi's Sarcoma cells:

### Cells and cell cultures

10 Human Kaposi's sarcoma cells lines KS 38 were derived from cutaneous biopsy of an AIDS patient as previously described (Lunardi-Iskandar, Y. et al. (1995) *J. Natl. Cancer. Inst.* 87, 974-981; Masood R. et al. (1994) *Human Retroviruses*. 10, 969-975). The cells were grown on 1.5% gelatin-coated flasks and were carried in RPMI 1640 with 15% fetal calf serum (FCS), 2mM glutamine, 1mM MEM Sodium  
15 Pyruvate, 0.05mM MEM Non-Essential Amino Acids, 1x MEM Amino Acids, 1% Nutridoma-HU (Boehringer Mannheim) and 50 mg/mL penicillin and 50 mg/ml streptomycin. Cultures were carried until near confluent prior to different treatments in the signaling studies described below. 293 cells were transfected with the FLT-4 gene and used as controls for detection of receptor protein as described (Lee, J. et al.  
20 (1996) *Proc. Natl. Acad. Sci. USA*. 93, 1988-1992).

### Reagents and antibodies

*RAFTK* antibodies were generated using GST fusion proteins to various domains of the molecule and immunizing New Zealand rabbits as previously  
25 described (Avraham S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751; Li J. et al. (1996) *Blood*. 88, 417-428). Using an ELISA assay, sera were screened for specific binding to *RAFTK*. Serum R4520 was chosen for further studies based on its titer in the ELISA. Serum R4520 did not cross react with FAK and was specific for *RAFTK*. Antibodies to the VEGF receptor FLK-1 and to the receptor FLT-4 were obtained  
30 from Genentech Inc. Antibodies to paxillin were obtained from Santa Cruz Biotechnology. Monoclonal anti-phosphotyrosine antibody (4G10) was obtained from Upstate Biotechnology Inc. (Lake Placid, NY). Electrophoresis reagents were obtained from Bio-Rad Laboratories (Hercules, CA). The Phorbol 12-myristate 13-acetate (PMA) and protease inhibitors leupeptin, aprotinin and alpha 1 antitrypsin and  
35 all other reagents were obtained from Sigma Co. (St.Louis, MO). The recombinant cytokines b-FGF, TNF- $\alpha$  and IL-6 were obtained from R&D systems. VRP, the ligand for the FLT-4 receptor, was obtained and expressed from a glioblastoma cell

line and purified as previously reported (Lee, J. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93, 1988-1992). Recombinant VEGF was obtained from Genentech, Inc. (South San Francisco, CA). Recombinant oncostatin M was obtained from the AIDS Reagent Bank (Bethesda, MD).

5

### Indirect immunofluorescence

KS 38 cells were cultured in Chamber Slides (Lab Tek) to 90% confluency. Cells were washed twice with cold PBS and then fixed for 30 min in 4% paraformaldehyde. Cells were washed 3X PBS and blocked for non-specific staining using 10% FCS in PBS for 30 minute on ice. FLT-4 and FLK-1 expression were determined using purified antiserum at a dilution 1:100 for one hour on ice. Normal rabbit serum was used as a control for non-specific staining. After washing cells 3X with PBS, cells were stained with secondary antibody conjugated to FITC (Boehringer Mannheim) at 1:500 dilution for one hour on ice. Proteins were visualized and photographed after washing 3X PBS using an inverted fluorescence microscope.

15

### Stimulation of cells

KS cells, grown to 80% confluence, were serum-starved for 16-18 hrs and washed twice with Hank's balanced salt solution (Gibco BRL) prior to PMA or cytokine treatments. KS cells were first treated with PMA to assess the effects of a chemical stimulus known to phosphorylate *RAFTK* in other cell systems (Avraham S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751). After a time course of stimulation with PMA was established, the effects of cytokines were studied. VEGF, VRP, TNF- $\alpha$ , oncostatin M, IL-6, or b-FGF were added to cultures at a range of concentrations for different time periods *in vitro*. After stimulation, cell lysates were directly prepared within the culture dish by lysis in 500  $\mu$ L modified RIPA (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 10 (g/ml of aprotinin, leupeptin and pepstatin, 10 mM sodium vanadate, 10 mM sodium fluoride and 10 mM sodium pyrophosphate) per dish at varying timepoints. Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by protein assay (Bio-Rad Laboratories).

25

30

### Immunoprecipitation and Western blot analysis

For immunoprecipitation studies, identical amounts of protein from each sample were clarified by incubation with protein A-Sepharose CL-4B (Pharmacia Biotech) for 1 h at 4°C. Following the removal of protein A-Sepharose by brief

35

centrifugation, the solution was incubated with different primary antibodies as detailed below for each experiment for 4 h or overnight at 4°C. Immunoprecipitations of the antibody-antigen complexes were performed by incubation for 2 h at 4°C with 75 µl of protein A-Sepharose (10% suspension). Non-specific bound proteins were removed by washing the Sepharose beads three times with modified RIPA buffer and one time with Phosphate buffered saline (PBS). Bound proteins were solubilized in 40 µl of 2 X Laemmli buffer and further analyzed by immunoblotting. Samples were separated on 7.5% SDS-PAGE and then transferred to nitrocellulose membranes. The membranes were blocked with 5% non-fat milk protein and probed with primary antibody for 2 h at RT or 4°C overnight. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) system (Amersham Corp., Arlington Heights, IL).

***Kaposi sarcoma cells express Flk-1 and Flt-4 receptors.***

To characterize the effects of different cytokines on KS cell signalling the KS 38 cell line was examined for expression of receptors for members of the VEGF family. Because KS spindle cells appear to be derived from lymphatic endothelium. Using indirect immunofluorescence (IFA), expression of both the FLK-1 receptor and the FLT-4 receptor was readily observed. The presence of both receptors was confirmed by Western blot using specific antisera to FLK-1 and FLT-4.

***RAFTK is expressed in Kaposi sarcoma cells and phosphorylated upon cytokine treatment.***

The KS cells were further characterized for expression of *RAFTK*. KS 38 cells expressed significant amounts of *RAFTK* protein as detected by Western blot and immunoprecipitation. Moreover, PMA treatment of KS 38 cells resulted in a time dependent phosphorylation of *RAFTK*.

Having established that *RAFTK* is expressed in KS 38 cells, whether or not treatment of these cells with VEGF or VRP, respective ligands for the FLK-1 and FLT-4 receptors, resulted in activation of signalling pathways that included *RAFTK* was investigated. There was a clear time dependent phosphorylation of *RAFTK* in response to VRP. Similar changes were observed following treatment with VEGF.

Previously, cytokines such as b-FGF, oncostatin M, IL-6 and TNF-α have been reported to promote *in vitro* proliferation of KS cells. The effects of treatment with these cytokines on *RAFTK* phosphorylation in KS 38 cells was analyzed. Each of these cytokines resulted in phosphorylation of this novel tyrosine kinase.

**Cytokine treatment of Kaposi sarcoma cells results in *RAFTK* association with the cytoskeletal protein paxillin.**

Following the observation that *RAFTK* was phosphorylated by cytokines like oncostatin M, IL-6, b-FGF and TNF- $\alpha$  known to stimulate KS cell growth, as well as the endothelial growth factors VEGF and VRP, whether this phosphorylation might modulate the association of *RAFTK* with certain cytoskeletal molecules was investigated. Using specific antibodies to *RAFTK* or paxillin, significantly increased association of these two molecules following cytokine treatments was found.

The development of authentic permanent KS cell lines has afforded the opportunity to characterize the surface structures of these cells and to examine which cytokines may modulate proliferation of the neoplasm. There is an extensive literature which supports a role for several cytokines in promoting KS cell growth via autocrine or paracrine mechanisms (See, e.g., Ba, G. et al. (1992) *J. Immunol.* 149, 3727-3734; Buonaguro, L. et al. (1992) *J. Virol.* 66, 7159-7167; Cai, J., et al. (1994) *Am. J. Pathol.* 145, 74-79). Characterization of signalling pathways in KS cells and the effects of these cytokines on such pathways have been less extensively explored. Amaral et. al found that OSM activated the MAP kinase pathway (Amaral M. C. et al. (1993) *J. Clin. Invest.* 92, 848-857) while Faris et. al reported that members of the Jak/Stat family of kinases known to participate in signalling via the gp 130 receptor were active in KS cells as well (Faris M. et al. (1996) *AIDS*. 10, 369-378). In these studies, the KS 38 cell line was derived from a patient with cutaneous KS as a model because of its previously characterized properties that closely correspond to those of primary pathological KS specimens (Lunardi-Iskandar, Y. et al. (1995) *J. Natl. Cancer. Inst.* 87, 974-981). The expression of novel receptors on KS 38 cells which are preferentially expressed in normal lymphatic endothelium was investigated, and characterization made of signalling pathways that may link surface receptor activation to the cytoskeleton in these cells.

The tyrosine kinase FLT-4 receptor is relatively restricted in expression in normal tissues, with prior studies indicating its presence on the surface of lymphatic endothelium (Kaipainen A. et al. (1995) *Proc. Natl. Acad. Sci. USA.* 92, 3566-3570; Kaipainen A. et al. (1993) *J. Exp. Med.* 178, 2077-2088; Pajusola K. et al. (1993) *Oncogene.* 8, 2931-2937). KS 38 cells express FLT-4 as well as the related FLK-1 receptor. The recently identified ligand VRP specifically binds to the FLT-4 receptor (Lee, J. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93, 1988-1992; Joukov V., et al. (1996) *EMBO J.* 15, 290-298), while FLK-1 is activated by VEGF. VRP, as well as VEGF, induced significant signalling changes in target KS 38 cells based on enhanced phosphorylation of proteins. Following this observation, investigation into signalling

molecules whose enhanced phosphorylation was a common response to these cytokines as well as those previously reported to stimulate KS cells was performed.

A variety of ligands and receptors of different molecular families have been implicated in the pathogenesis of KS. To survey this range of cytokines, 5 representative cytokines from each family was chosen and a comparative analysis was made of b-FGF, TNF- $\alpha$ , oncostatin M, and IL-6 with VEGF and VRP. They all signalled via the recently identified *RAFTK* molecule.

*RAFTK* functions as a "platform kinase" upon which a number of intracytoplasmic kinases and adaptor molecules converge. The convergence of such 10 molecules facilitates transmission of surface signals to the cytoskeleton. In this study, the phosphorylation of *RAFTK* by a variety of cytokines which belong to distinctly different families is described. Specifically, b-FGF, VEGF and VRP signal through receptor tyrosine kinases, OSM and IL-6 bind to a dual receptor with a specific chain and a shared gp130 chain, and TNF- $\alpha$  binds to the CD95 receptor family linked to 15 apoptosis. It appears, from this KS cell line model, that *RAFTK* may participate in each of these diverse receptor activated pathways. Based on the observations reported here, *RAFTK* functions in KS cells to transduce receptor signals via association with cytoskeletal molecules such as paxillin. *RAFTK* activation likely participates in a final common pathway for KS cell growth. Given the accessibility of cutaneous KS 20 lesions to locally applied treatments, specific inhibitors of *RAFTK* are particularly useful in treatment of this disorder.

### Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more 25 than routine experimentation, many equivalents of the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

5

(i) APPLICANTS: Shalom Avraham et al.

(ii) TITLE OF INVENTION: Novel RAFTK Signaling Molecules and Uses Therefor

10

(iii) NUMBER OF SEQUENCES: 18

(iv) CORRESPONDENCE ADDRESS:

15

(A) ADDRESSEE: LAHIVE & COCKFIELD  
(B) STREET: 60 State Street, suite 510  
(C) CITY: Boston  
(D) STATE: Massachusetts  
(E) COUNTRY: USA  
(F) ZIP: 02109-1875

20

(v) COMPUTER READABLE FORM:

25

(A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

30

(A) APPLICATION NUMBER:  
(B) FILING DATE: August 23, 1996  
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

35

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Jean M. Silveri  
(B) REGISTRATION NUMBER: 39,030  
(C) REFERENCE/DOCKET NUMBER: NER-255

40

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617)227-7400  
(B) TELEFAX: (617)227-5941

45

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

50

(A) LENGTH: 3621 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

55

(ix) FEATURE:



(A) NAME/KEY: CDS

(B) LOCATION: 294..3321

5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
	AACCACAAGT CAAATAGAAA GAAGTTAAAA GAATGTTTAT GCAAACACAT GAGAAAAGAA	60
	GGGTGCAGAT GAGAATAGGG GTGTGGTTAA CAACTCAGAG GAGGAGGGAG AATCTAACCT	120
10	GTCAGCCCTT TTA CTAGCC ACAGCCTCCG GAGCCGTTGC ACACCTACCT GCCCGGCCGA	180
	CTTACCTGTA CTTGCCGCCG TCCCGGCTCA CCTGGCGGTG CCCGAGGAGT AGTCGCTGGA	240
15	GTCCGCGCCT CCCTGGGACT GCAATGTGCC GATCTTAGCT GCTGCCTGAG AGG ATG	296
	Met	
	1	
	TCT GGG GTG TCC GAG CCC CTG AGT CGA GTA AAG TTG GGC ACG TTA CGC	344
20	Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Leu Gly Thr Leu Arg	
	5 10 15	
	CGG CCT GAA GGC CCT GCA GAG CCC ATG GTG GTG GTA CCA GTA GAT GTG	392
25	Arg Pro Glu Gly Pro Ala Glu Pro Met Val Val Val Pro Val Asp Val	
	20 25 30	
	GAA AAG GAG GAC GTG CGT ATC CTC AAG GTC TGC TTC TAT AGC AAC AGC	440
	Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn Ser	
30	35 40 45	
	TTC AAT CCT GGG AAA AAC TTC AAA CTG GTC AAA TGC ACT GTC CAG ACG	488
	Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln Thr	
	50 55 60 65	
35	GAG ATC CGG GAG ATC ATC ACC TCC ATC CTG CTG AGC GGG CGG ATC GGG	536
	Glu Ile Arg Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile Gly	
	70 75 80	
	CCC AAC ATC CGG TTG GCT GAG TGC TAT GGG CTG AGG CTG AAG CAC ATG	584
40	Pro Asn Ile Arg Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His Met	
	85 90 95	
	AAG TCC GAT GAG ATC CAC TGG CTG CAC CCA CAG ATG ACG GTG GGT GAG	632
45	Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly Glu	
	100 105 110	
	GTG CAG GAC AAG TAT GAG TGT CTG CAC GTG GAA GCC GAG TGG AGG TAT	680
	Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg Tyr	
50	115 120 125	
	GAC CTT CAA ATC CGC TAC TTG CCA GAA GAC TTC ATG GAG AGC CTG AAG	728
	Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu Lys	
	130 135 140 145	
55	GAG GAC AGG ACC ACG CTG CTC TAT TTT TAC CAA CAG CTC CGG AAC GAC	776
	Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn Asp	

					150					155					160		
5	TAC	ATG	CAG	CGC	TAC	GCC	AGC	AAG	GTC	AGC	GAG	GGC	ATG	GCC	CTG	CAG	824
	Tyr	Met	Gln	Arg	Tyr	Ala	Ser	Lys	Val	Ser	Glu	Gly	Met	Ala	Leu	Gln	
				165					170					175			
10	CTG	GGC	TGC	CTG	GAG	CTC	AGG	CGG	TTC	TTC	AAG	GAT	ATG	CCC	CAC	AAT	872
	Leu	Gly	Cys	Leu	Glu	Leu	Arg	Arg	Phe	Phe	Lys	Asp	Met	Pro	His	Asn	
			180					185					190				
	GCA	CTT	GAC	AAG	AAG	TCC	AAC	TTC	GAG	CTC	CTA	GAA	AAG	GAA	GTG	GGG	920
	Ala	Leu	Asp	Lys	Lys	Ser	Asn	Phe	Glu	Leu	Leu	Glu	Lys	Glu	Val	Gly	
		195					200					205					
15	CTG	GAC	TTG	TTT	TTC	CCA	AAG	CAG	ATG	CAG	GAG	AAC	TTA	AAG	CCC	AAA	968
	Leu	Asp	Leu	Phe	Phe	Pro	Lys	Gln	Met	Gln	Glu	Asn	Leu	Lys	Pro	Lys	
		210				215					220					225	
20	CAG	TTC	CGG	AAG	ATG	ATC	CAG	CAG	ACC	TTC	CAG	CAG	TAC	GCC	TCG	CTC	1016
	Gln	Phe	Arg	Lys	Met	Ile	Gln	Gln	Thr	Phe	Gln	Gln	Tyr	Ala	Ser	Leu	
					230					235					240		
25	AGG	GAG	GAG	GAG	TGC	GTC	ATG	AAG	TTC	TTC	AAC	ACT	CTC	GCC	GGC	TTC	1064
	Arg	Glu	Glu	Glu	Cys	Val	Met	Lys	Phe	Phe	Asn	Thr	Leu	Ala	Gly	Phe	
				245					250					255			
30	GCC	AAC	ATC	GAC	CAG	GAG	ACC	TAC	CGC	TGT	GAA	CTC	ATT	CAA	GGA	TGG	1112
	Ala	Asn	Ile	Asp	Gln	Glu	Thr	Tyr	Arg	Cys	Glu	Leu	Ile	Gln	Gly	Trp	
			260					265					270				
	AAC	ATT	ACT	GTG	GAC	CTG	GTC	ATT	GGC	CCT	AAA	GGG	ATC	CGC	CAG	CTG	1160
	Asn	Ile	Thr	Val	Asp	Leu	Val	Ile	Gly	Pro	Lys	Gly	Ile	Arg	Gln	Leu	
		275					280					285					
35	ACT	AGT	CAG	GAC	GCA	AAG	CCC	ACC	TGC	CTG	GCC	GAG	TTC	AAG	CAG	ATC	1208
	Thr	Ser	Gln	Asp	Ala	Lys	Pro	Thr	Cys	Leu	Ala	Glu	Phe	Lys	Gln	Ile	
		290				295					300					305	
40	AGG	TCC	ATC	AGG	TGC	CTC	CCG	CTG	GAG	GAG	GGC	CAG	GCA	GTA	CTT	CAG	1256
	Arg	Ser	Ile	Arg	Cys	Leu	Pro	Leu	Glu	Glu	Gly	Gln	Ala	Val	Leu	Gln	
					310					315					320		
45	CTG	GGC	ATT	GAA	GGT	GCC	CCC	CAG	GCC	TTG	TCC	ATC	AAA	ACC	TCA	TCC	1304
	Leu	Gly	Ile	Glu	Gly	Ala	Pro	Gln	Ala	Leu	Ser	Ile	Lys	Thr	Ser	Ser	
				325				330						335			
50	CTA	GCA	GAG	GCT	GAG	AAC	ATG	GCT	GAC	CTC	ATA	GAC	GGC	TAC	TGC	CGG	1352
	Leu	Ala	Glu	Ala	Glu	Asn	Met	Ala	Asp	Leu	Ile	Asp	Gly	Tyr	Cys	Arg	
			340				345						350				
	CTG	CAG	GGT	GAG	CAC	CAA	GGC	TCT	CTC	ATC	ATC	CAT	CCT	AGG	AAA	GAT	1400
	Leu	Gln	Gly	Glu	His	Gln	Gly	Ser	Leu	Ile	Ile	His	Pro	Arg	Lys	Asp	
		355					360				365						
55	GGT	GAG	AAG	CGG	AAC	AGC	CTG	CCC	CAG	ATC	CCC	ATG	CTA	AAC	CTG	GAG	1448
	Gly	Glu	Lys	Arg	Asn	Ser	Leu	Pro	Gln	Ile	Pro	Met	Leu	Asn	Leu	Glu	

	370				375					380				385			
	GCC	CGG	CGG	TCC	CAC	CTC	TCA	GAG	AGC	TGC	AGC	ATA	GAG	TCA	GAC	ATC	1496
5	Ala	Arg	Arg	Ser	His	Leu	Ser	Glu	Ser	Cys	Ser	Ile	Glu	Ser	Asp	Ile	
					390					395					400		
	TAC	GCA	GAG	ATT	CCC	GAC	GAA	ACC	CTG	CGA	AGG	CCC	GGA	GGT	CCA	CAG	1544
	Tyr	Ala	Glu	Ile	Pro	Asp	Glu	Thr	Leu	Arg	Arg	Pro	Gly	Gly	Pro	Gln	
				405					410					415			
10	TAT	GGC	ATT	GCC	CGT	GAA	GAT	GTG	GTC	CTG	AAT	CGT	ATT	CTT	GGG	GAA	1592
	Tyr	Gly	Ile	Ala	Arg	Glu	Asp	Val	Val	Leu	Asn	Arg	Ile	Leu	Gly	Glu	
			420					425					430				
15	GGC	TTT	TTT	GGG	GAG	GTC	TAT	GAA	GGT	GTC	TAC	ACA	AAT	CAT	AAA	GGG	1640
	Gly	Phe	Phe	Gly	Glu	Val	Tyr	Glu	Gly	Val	Tyr	Thr	Asn	His	Lys	Gly	
		435					440					445					
	GAG	AAA	ATC	AAT	GTA	GCT	GTC	AAG	ACC	TGC	AAG	AAA	GAC	TGC	ACT	CTG	1688
20	Glu	Lys	Ile	Asn	Val	Ala	Val	Lys	Thr	Cys	Lys	Lys	Asp	Cys	Thr	Leu	
		450				455					460					465	
	GAC	AAC	AAG	GAG	AAG	TTC	ATG	AGC	GAG	GCA	GTG	ATC	ATG	AAG	AAC	CTC	1736
25	Asp	Asn	Lys	Glu	Lys	Phe	Met	Ser	Glu	Ala	Val	Ile	Met	Lys	Asn	Leu	
					470					475					480		
	GAC	CAC	CCG	CAC	ATC	GTG	AAG	CTG	ATC	GGC	ATC	ATT	GAA	GAG	GAG	CCC	1784
	Asp	His	Pro	His	Ile	Val	Lys	Leu	Ile	Gly	Ile	Ile	Glu	Glu	Glu	Pro	
				485					490					495			
30	ACC	TGG	ATC	ATC	ATG	GAA	TTG	TAT	CCC	TAT	GGG	GAG	CTG	GGC	CAC	TAC	1832
	Thr	Trp	Ile	Ile	Met	Glu	Leu	Tyr	Pro	Tyr	Gly	Glu	Leu	Gly	His	Tyr	
			500					505					510				
35	CTG	GAG	CGG	AAC	AAG	AAC	TCC	CTG	AAG	GTG	CTC	ACC	CTC	GTG	CTG	TAC	1880
	Leu	Glu	Arg	Asn	Lys	Asn	Ser	Leu	Lys	Val	Leu	Thr	Leu	Val	Leu	Tyr	
		515					520					525					
	TCA	CTG	CAG	ATA	TGC	AAA	GCC	ATG	GCC	TAC	CTG	GAG	AGC	ATC	AAC	TGC	1928
40	Ser	Leu	Gln	Ile	Cys	Lys	Ala	Met	Ala	Tyr	Leu	Glu	Ser	Ile	Asn	Cys	
		530				535					540					545	
	GTG	CAC	AGG	GAC	ATT	GCT	GTC	CGG	AAC	ATC	CTG	GTG	GCC	TCC	CCT	GAG	1976
45	Val	His	Arg	Asp	Ile	Ala	Val	Arg	Asn	Ile	Leu	Val	Ala	Ser	Pro	Glu	
					550					555					560		
	TGT	GTG	AAG	CTG	GGG	GAC	TTT	GGT	CTT	TCC	CGG	TAC	ATT	GAG	GAC	GAG	2024
	Cys	Val	Lys	Leu	Gly	Asp	Phe	Gly	Leu	Ser	Arg	Tyr	Ile	Glu	Asp	Glu	
				565					570					575			
50	GAC	TAT	TAC	AAA	GCC	TCT	GTG	ACT	CGT	CTC	CCC	ATC	AAA	TGG	ATG	TCC	2072
	Asp	Tyr	Tyr	Lys	Ala	Ser	Val	Thr	Arg	Leu	Pro	Ile	Lys	Trp	Met	Ser	
				580					585				590				
55	CCA	GAG	TCC	ATT	AAC	TTC	CGA	CGC	TTC	ACG	ACA	GCC	AGT	GAC	GTC	TGG	2120
	Pro	Glu	Ser	Ile	Asn	Phe	Arg	Arg	Phe	Thr	Thr	Ala	Ser	Asp	Val	Trp	

	595	600	605	
5	ATG TTC GCC GTG TGC ATG TGG GAG ATC CTG AGC TTT GGG AAG CAG CCC Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln Pro 610 615 620 625	2168		
10	TTC TTC TGG CTG GAG AAC AAG GAT GTC ATC GGG GTG CTG GAG AAA GGA Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys Gly 630 635 640	2216		
15	GAC CGG CTG CCC AAG CCT GAT CTC TGT CCA CCG GTC CTT TAT ACC CTC Asp Arg Leu Pro Lys Pro Asp Leu Cys Pro Pro Val Leu Tyr Thr Leu 645 650 655	2264		
20	ATG ACC CGC TGC TGG GAC TAC GAC CCC AGT GAC CGG CCC CGC TTC ACC Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe Thr 660 665 670	2312		
25	GAG CTG GTG TGC AGC CTC AGT GAC GTT TAT CAG ATG GAG AAG GAC ATT Glu Leu Val Cys Ser Leu Ser Asp Val Tyr Gln Met Glu Lys Asp Ile 675 680 685	2360		
30	GCC ATG GAG CAA GAG AGG AAT GCT CGC TAC CGA ACC CCC AAA ATC TTG Ala Met Glu Gln Glu Arg Asn Ala Arg Tyr Arg Thr Pro Lys Ile Leu 690 695 700 705	2408		
35	GAG CCC ACA GCC TTC CAG GAA CCC CCA CCC AAG CCC AGC CGA CCT AAG Glu Pro Thr Ala Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro Lys 710 715 720	2456		
40	TAC AGA CCC CCT CCG CAA ACC AAC CTC CTG GCT CCA AAG CTG CAG TTC Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln Phe 725 730 735	2504		
45	CAG GTT CCT GAG GGT CTG TGT GCC AGC TCT CCT ACG CTC ACC AGC CCT Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser Pro 740 745 750	2552		
50	ATG GAG TAT CCA TCT CCC GTT AAC TCA CTG CAC ACC CCA CCT CTC CAC Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu His 755 760 765	2600		
55	CGG CAC AAT GTC TTC AAA CGC CAC AGC ATG CGG GAG GAG GAC TTC ATC Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe Ile 770 775 780 785	2648		
60	CAA CCC AGC AGC CGA GAA GAG GCC CAG CAG CTG TGG GAG GCT GAA AAG Gln Pro Ser Ser Arg Glu Glu Ala Gln Leu Trp Glu Ala Glu Lys 790 795 800	2696		
65	GTC AAA ATG CGG CAA ATC CTG GAC AAA CAG CAG AAG CAG ATG GTG GAG Val Lys Met Arg Gln Ile Leu Asp Lys Gln Gln Lys Gln Met Val Glu 805 810 815	2744		
70	GAC TAC CAG TGG CTC AGG CAG GAG GAG AAG TCC CTG GAC CCC ATG GTT Asp Tyr Gln Trp Leu Arg Gln Glu Glu Lys Ser Leu Asp Pro Met Val	2792		

	820					825					830							
5	TAT	ATG	AAT	GAT	AAG	TCC	CCA	TTG	ACG	CCA	GAG	AAG	GAG	GTC	GGC	TAC	2840	
	Tyr	Met	Asn	Asp	Lys	Ser	Pro	Leu	Thr	Pro	Glu	Lys	Glu	Val	Gly	Tyr		
	835					840					845							
10	CTG	GAG	TTC	ACA	GGG	CCC	CCA	CAG	AAG	CCC	CCG	AGG	CTG	GGC	GCA	CAG	2888	
	Leu	Glu	Phe	Thr	Gly	Pro	Pro	Gln	Lys	Pro	Pro	Arg	Leu	Gly	Ala	Gln		
	850					855					860					865		
15	TCC	ATC	CAG	CCC	ACA	GCT	AAC	CTG	GAC	CGG	ACC	GAT	GAC	CTG	GTG	TAC	2936	
	Ser	Ile	Gln	Pro	Thr	Ala	Asn	Leu	Asp	Arg	Thr	Asp	Asp	Leu	Val	Tyr		
	870					875					880							
20	CTC	AAT	GTC	ATG	GAG	CTG	GTG	CGG	GCC	GTG	CTG	GAG	CTC	AAG	AAT	GAG	2984	
	Leu	Asn	Val	Met	Glu	Leu	Val	Arg	Ala	Val	Leu	Glu	Leu	Lys	Asn	Glu		
	885					890					895							
25	CTC	TGT	CAG	CTG	CCC	CCC	GAG	GGC	TAC	GTG	GTG	GTG	GTG	AAG	AAT	GTG	3032	
	Leu	Cys	Gln	Leu	Pro	Pro	Glu	Gly	Tyr	Val	Val	Val	Val	Lys	Asn	Val		
	900					905					910							
30	GGG	CTG	ACC	CTG	CGG	AAG	CTC	ATC	GGG	AGC	GTG	GAT	GAT	CTC	CTG	CCT	3080	
	Gly	Leu	Thr	Leu	Arg	Lys	Leu	Ile	Gly	Ser	Val	Asp	Asp	Leu	Leu	Pro		
	915					920					925							
35	TCC	TTG	CCG	TCA	TCT	TCA	CGG	ACA	GAG	ATC	GAG	GGC	ACC	CAG	AAA	CTG	3128	
	Ser	Leu	Pro	Ser	Ser	Ser	Arg	Thr	Glu	Ile	Glu	Gly	Thr	Gln	Lys	Leu		
	930					935					940					945		
40	CTC	AAC	AAA	GAC	CTG	GCA	GAG	CTC	ATC	AAC	AAG	ATG	CGG	CTG	GCG	CAG	3176	
	Leu	Asn	Lys	Asp	Leu	Ala	Glu	Leu	Ile	Asn	Lys	Met	Arg	Leu	Ala	Gln		
	950					955					960							
45	CAG	AAC	GCC	GTG	ACC	TCC	CTG	AGT	GAG	GAG	TGC	AAG	AGG	CAG	ATG	CTG	3224	
	Gln	Asn	Ala	Val	Thr	Ser	Leu	Ser	Glu	Glu	Cys	Lys	Arg	Gln	Met	Leu		
	965					970					975							
50	ACG	GCT	TCA	CAC	ACC	CTG	GCT	GTG	GAC	GCC	AAG	AAC	CTG	CTC	GAC	GCT	3272	
	Thr	Ala	Ser	His	Thr	Leu	Ala	Val	Asp	Ala	Lys	Asn	Leu	Leu	Asp	Ala		
	980					985					990							
55	GTG	GAC	CAG	GCC	AAG	GTT	CTG	GCC	AAT	CTG	GCC	CAC	CCA	CCT	GCA	GAG	T	3321
	Val	Asp	Gln	Ala	Lys	Val	Leu	Ala	Asn	Leu	Ala	His	Pro	Pro	Ala	Glu		
	995					1000					1005							
	GACGGAGGGT GGGGGCCACC TGCCTGCGTC TTCCGCCCTT GCCTGCCATG TACCTCCCCT																3381	
	GCCTTGCTGT TGGTCATGTG GGTCTTCCAG GGAGAAGGCC AAGGGGAGTC ACCTTCCCTT																3441	
	GCCACTTTGC ACGACGCCCT CTCCCCACCC CTACCCCTGG CTGTACTGCT CAGGCTGCAG																3501	
	CTGGACAGAG GGGACTCTGG GCTATGGACA CAGGGTGACG GTGACAAAGA TGGCTCAGAG																3561	
	GGGGACTGCT GCTGCCTGGC CACTGCTCCC TAAGCCAGCC TGGTCCATGC AGGGGGCTCG																3621	

## (2) INFORMATION FOR SEQ ID NO:2:

## 5 (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1009 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

## 10 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

15 Met Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Leu Gly Thr Leu
    1           5           10           15

Arg Arg Pro Glu Gly Pro Ala Glu Pro Met Val Val Val Pro Val Asp
    20           25           30

20 Val Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn
    35           40           45

Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln
    50           55           60

25 Thr Glu Ile Arg Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile
    65           70           75           80

Gly Pro Asn Ile Arg Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His
    85           90           95

30 Met Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly
    100          105          110

35 Glu Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg
    115          120          125

Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu
    130          135          140

40 Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn
    145          150          155          160

Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val Ser Glu Gly Met Ala Leu
    165          170          175

Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe Phe Lys Asp Met Pro His
    180          185          190

50 Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu Leu Leu Glu Lys Glu Val
    195          200          205

Gly Leu Asp Leu Phe Phe Pro Lys Gln Met Gln Glu Asn Leu Lys Pro
    210          215          220

55 Lys Gln Phe Arg Lys Met Ile Gln Gln Thr Phe Gln Gln Tyr Ala Ser

```

- 137 -

	225		230		235		240
	Leu Arg Glu Glu Glu Cys Val Met Lys Phe Phe Asn Thr Leu Ala Gly						
		245		250		255	
5	Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg Cys Glu Leu Ile Gln Gly						
		260		265		270	
10	Trp Asn Ile Thr Val Asp Leu Val Ile Gly Pro Lys Gly Ile Arg Gln						
		275		280		285	
	Leu Thr Ser Gln Asp Ala Lys Pro Thr Cys Leu Ala Glu Phe Lys Gln						
		290		295		300	
15	Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu Glu Gly Gln Ala Val Leu						
		305		310		315	
	Gln Leu Gly Ile Glu Gly Ala Pro Gln Ala Leu Ser Ile Lys Thr Ser						
		325		330		335	
20	Ser Leu Ala Glu Ala Glu Asn Met Ala Asp Leu Ile Asp Gly Tyr Cys						
		340		345		350	
25	Arg Leu Gln Gly Glu His Gln Gly Ser Leu Ile Ile His Pro Arg Lys						
		355		360		365	
	Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln Ile Pro Met Leu Asn Leu						
		370		375		380	
30	Glu Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp						
		385		390		395	
	Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro						
		405		410		415	
35	Gln Tyr Gly Ile Ala Arg Glu Asp Val Val Leu Asn Arg Ile Leu Gly						
		420		425		430	
40	Glu Gly Phe Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys						
		435		440		445	
	Gly Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr						
		450		455		460	
45	Leu Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn						
		465		470		475	
	Leu Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu						
		485		490		495	
50	Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His						
		500		505		510	
55	Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Leu Thr Leu Val Leu						
		515		520		525	

- 138 -

Tyr Ser Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn  
 530 535 540

5 Cys Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro  
 545 550 555 560

Glu Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp  
 565 570 575

10 Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met  
 580 585 590

Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val  
 595 600 605

15 Trp Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln  
 610 615 620

20 Pro Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys  
 625 630 635 640

Gly Asp Arg Leu Pro Lys Pro Asp Leu Cys Pro Pro Val Leu Tyr Thr  
 645 650 655

25 Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe  
 660 665 670

Thr Glu Leu Val Cys Ser Leu Ser Asp Val Tyr Gln Met Glu Lys Asp  
 675 680 685

30 Ile Ala Met Glu Gln Glu Arg Asn Ala Arg Tyr Arg Thr Pro Lys Ile  
 690 695 700

35 Leu Glu Pro Thr Ala Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro  
 705 710 715 720

Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln  
 725 730 735

40 Phe Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser  
 740 745 750

Pro Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu  
 755 760 765

45 His Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe  
 770 775 780

50 Ile Gln Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu  
 785 790 795 800

Lys Val Lys Met Arg Gln Ile Leu Asp Lys Gln Gln Lys Gln Met Val  
 805 810 815

55 Glu Asp Tyr Gln Trp Leu Arg Gln Glu Glu Lys Ser Leu Asp Pro Met  
 820 825 830



[illegible]

(2) INFORMATION FOR SEQ ID NO:3:

40

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4029 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

45

## (ii) MOLECULE TYPE: cDNA

50

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 202..3229

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	GAGAGCAGCA GGGGTGTGGT TAACGACCGA GAGGAGGAGG GGGAAAAACA ACCTGTCAGC	60
	CTCTTACTCA GCCTCTGCAG GCAGAGCCGC GCGTCCTACC TGCGGCGGCT GCGCTCACCT	120
5	GGCCCAGCCC GGAGCCCTGG CCCGAGTCCG CGCCTCGCCC GAGGGACTGC AATGTGCCGG	180
	TCCTAGCTGC AGTCTGAGAG G ATG TCC GGG GTG TCT GAG CCC TTG AGC CGT	231
	Met Ser Gly Val Ser Glu Pro Leu Ser Arg	
	1 5 10	
10	GTA AAA GTG GGC ACT TTA CGC CGG CCT GAG GGC CCC CCA GAG CCC ATG	279
	Val Lys Val Gly Thr Leu Arg Arg Pro Glu Gly Pro Pro Glu Pro Met	
	15 20 25	
15	GTG GTG GTA CCA GTG GAT GTG GAG AAG GAA GAC GTG CGC ATC CTC AAG	327
	Val Val Val Pro Val Asp Val Glu Lys Glu Asp Val Arg Ile Leu Lys	
	30 35 40	
20	GTC TGC TTC TAC AGC AAC AGC TTC AAC CCA GGG AAG AAC TTC AAG CTT	375
	Val Cys Phe Tyr Ser Asn Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu	
	45 50 55	
25	GTC AAA TGC ACA GTG CAG ACA GAG ATC CAG GAG ATC ATC ACC TCC ATC	423
	Val Lys Cys Thr Val Gln Thr Glu Ile Gln Glu Ile Ile Thr Ser Ile	
	60 65 70	
30	CTC CTG AGT GGG CGA ATA GGG CCC AAC ATC CAG CTG GCT GAA TGC TAT	471
	Leu Leu Ser Gly Arg Ile Gly Pro Asn Ile Gln Leu Ala Glu Cys Tyr	
	75 80 85 90	
35	GGG CTG AGG CTG AAG CAC ATG AAG TCA GAC GAG ATC CAC TGG CTG CAC	519
	Gly Leu Arg Leu Lys His Met Lys Ser Asp Glu Ile His Trp Leu His	
	95 100 105	
40	CCA CAG ATG ACC GTG GGC GAA GTG CAG GAC AAG TAT GAA TGT CTA CAC	567
	Pro Gln Met Thr Val Gly Glu Val Gln Asp Lys Tyr Glu Cys Leu His	
	110 115 120	
45	GTG GAA GCT GAG TGG AGG TAT GAC CTT CAA ATC CGC TAC TTG CCG GAA	615
	Val Glu Ala Glu Trp Arg Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu	
	125 130 135	
50	GAC TTC ATG GAG AGC CTG AAA GAA GAC AGG ACC ACA TTG CTG TAC TTT	663
	Asp Phe Met Glu Ser Leu Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe	
	140 145 150	
55	TAT CAA CAG CTC CGG AAT GAC TAC ATG CAA CGC TAC GCC AGC AAG GTC	711
	Tyr Gln Gln Leu Arg Asn Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val	
	155 160 165 170	
55	AGT GAA GGC ATG GCT CTG CAG CTG GGC TGT CTG GAG CTC AGG AGA TTC	759
	Ser Glu Gly Met Ala Leu Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe	
	175 180 185	
55	TTC AAG GAC ATG CCC CAC AAT GCA CTG GAC AAA AAG TCC AAC TTT GAA	807
	Phe Lys Asp Met Pro His Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu	

	190	195	200	
5	CTC CTG GAA AAA GAA GTC GGT CTG GAC CTG TTT TTC CCA AAG CAG ATG Leu Leu Glu Lys Glu Val Gly Leu Asp Leu Phe Phe Pro Lys Gln Met 205 210 215	855		
10	CAG GAA AAC TTA AAG CCC AAG CAG TTC CGG AAG ATG ATC CAG CAG ACC Gln Glu Asn Leu Lys Pro Lys Gln Phe Arg Lys Met Ile Gln Gln Thr 220 225 230	903		
15	TTC CAG CAG TAT GCA TCA CTC CGG GAG GAA GAG TGT GTC ATG AAA TTC Phe Gln Gln Tyr Ala Ser Leu Arg Glu Glu Cys Val Met Lys Phe 235 240 245 250	951		
20	TTC AAT ACC CTA GCG GGC TTT GCC AAC ATT GAC CAG GAG ACC TAC CGC Phe Asn Thr Leu Ala Gly Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg 255 260 265	999		
25	TGC GAA CTC ATT CAA GGA TGG AAC ATT ACT GTG GAC CTG GTC ATC GGC Cys Glu Leu Ile Gln Gly Trp Asn Ile Thr Val Asp Leu Val Ile Gly 270 275 280	1047		
30	CCT AAA GGC ATC CGT CAG CTG ACA AGT CAA GAT ACA AAG CCC ACC TGC Pro Lys Gly Ile Arg Gln Leu Thr Ser Gln Asp Thr Lys Pro Thr Cys 285 290 295	1095		
35	CTG GCC GAG TTT AAG CAG ATC AGA TCC ATC AGG TGC CTC CCA TTG GAA Leu Ala Glu Phe Lys Gln Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu 300 305 310	1143		
40	GAG ACC CAG GCA GTC CTG CAG CTG GGC ATC GAG GGT GCC CCC CAG TCC Glu Thr Gln Ala Val Leu Gln Leu Gly Ile Glu Gly Ala Pro Gln Ser 315 320 325 330	1191		
45	TTG TCT ATC AAA ACG TCG TCC CTG GCA GAG GCT GAG AAC ATG GCT GAT Leu Ser Ile Lys Thr Ser Ser Leu Ala Glu Ala Glu Asn Met Ala Asp 335 340 345	1239		
50	CTC ATA GAT GGC TAC TGC AGG CTG CAA GGA GAA CAT AAG GGC TCT CTC Leu Ile Asp Gly Tyr Cys Arg Leu Gln Gly Glu His Lys Gly Ser Leu 350 355 360	1287		
55	ATC ATG CAT GCC AAG AAA GAT GGT GAG AAG AGG AAC AGC CTG CCT CAG Ile Met His Ala Lys Lys Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln 365 370 375	1335		
60	ATC CCC ACA CTA AAC CTG GAG GCT CGG CGG TCG CAC CTC TCA GAA AGC Ile Pro Thr Leu Asn Leu Glu Ala Arg Arg Ser His Leu Ser Glu Ser 380 385 390	1383		
65	TGC AGC ATA GCG TCA GAC ATC TAT GCG GAG ATT CCC GAT GAG ACC CTG Cys Ser Ile Glu Ser Asp Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu 395 400 405 410	1431		
70	CGA AGA CCA GGA GGT CCA CAG TAC GGT GTT GCC CGT GAA GAA GTA GTT Arg Arg Pro Gly Gly Pro Gln Tyr Gly Val Ala Arg Glu Glu Val Val	1479		

	415	420	425	
5	CTT AAC CGC ATT CTG GGT GAA GGC TTC TTT GGG GAG GTC TAT GAA GGT Leu Asn Arg Ile Leu Gly Glu Gly Phe Gly Glu Val Tyr Glu Gly 430 435 440			1527
10	GTC TAC ACG AAC CAC AAA GGG GAA AAA ATT AAT GTG GCC GTC AAG ACC Val Tyr Thr Asn His Lys Gly Glu Lys Ile Asn Val Ala Val Lys Thr 445 450 455			1575
15	TGT AAG AAA GAC TGT ACC CAG GAC AAC AAG GAG AAG TTC ATG AGT GAG Cys Lys Lys Asp Cys Thr Gln Asp Asn Lys Glu Lys Phe Met Ser Glu 460 465 470			1623
20	GCA GTG ATC ATG AAG AAT CTT GAC CAC CCT CAC ATC GTG AAG CTG ATT Ala Val Ile Met Lys Asn Leu Asp His Pro His Ile Val Lys Leu Ile 475 480 485 490			1671
25	GGC ATC ATT GAA GAG GAA CCC ACC TGG ATT ATC ATG GAA CTG TAT CCT Gly Ile Ile Glu Glu Glu Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro 495 500 505			1719
30	TAT GGG GAG CTG GGA CAC TAC CTG GAA CGA AAT AAA AAC TCC CTG AAG Tyr Gly Glu Leu Gly His Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys 510 515 520			1767
35	GTA CCC ACT CTG GTC CTG TAC ACC CTA CAG ATA TGC AAA GCC ATG GCC Val Pro Thr Leu Val Leu Tyr Thr Leu Gln Ile Cys Lys Ala Met Ala 525 530 535			1815
40	TAT CTG GAG AGC ATC AAC TGT GTG CAC AGG GAT ATT GCT GTC CGG AAC Tyr Leu Glu Ser Ile Asn Cys Val His Arg Asp Ile Ala Val Arg Asn 540 545 550			1863
45	ATC CTG GTG GCC TCT CCT GAG TGT GTG AAG CTG GGG GAC TTT GGG CTC Ile Leu Val Ala Ser Pro Glu Cys Val Lys Leu Gly Asp Phe Gly Leu 555 560 565 570			1911
50	TCC CGG TAC ATT GAG GAC GAA GAC TAT TAC AAA GCC TCT GTG ACA CGT Ser Arg Tyr Ile Glu Asp Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg 575 580 585			1959
55	CTA CCC ATC AAA TGG ATG TCC CCC GAG TCC ATC AAC TTC CGC CGC TTC Leu Pro Ile Lys Trp Met Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe 590 595 600			2007
60	ACA ACC GCC AGT GAT GTC TGG ATG TTT GCT GTA TGC ATG TGG GAG ATC Thr Thr Ala Ser Asp Val Trp Met Phe Ala Val Cys Met Trp Glu Ile 605 610 615			2055
65	CTC AGC TTT GGG AAG CAG CCT TTC TTC TGG CTC GAA AAT AAG GAT GTC Leu Ser Phe Gly Lys Gln Pro Phe Phe Trp Leu Glu Asn Lys Asp Val 620 625 630			2103
70	ATC GGA GTG CTG GAG AAA GGG GAC AGG CTG CCC AAG CCC GAA CTC TGT Ile Gly Val Leu Glu Lys Gly Asp Arg Leu Pro Lys Pro Glu Leu Cys 635 640 645 650			2151

	635		640		645		650	
5	CCG CCT GTC CTT TAC ACA CTC ATG ACT CGC TGC TGG GAC TAC GAC CCC Pro Pro Val Leu Tyr Thr Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro		655		660		665	2199
10	AGT GAC CGG CCC CGC TTC ACG GAG CTT GTG TGC AGC CTC AGT GAC ATT Ser Asp Arg Pro Arg Phe Thr Glu Leu Val Cys Ser Leu Ser Asp Ile		670		675		680	2247
15	TAT CAG ATG GAG AAG GAC ATT GCC ATA GAG CAA GAA AGG AAT GCT CGC Tyr Gln Met Glu Lys Asp Ile Ala Ile Glu Gln Glu Arg Asn Ala Arg		685		690		695	2295
20	TAC CGA CCC CCT AAA ATA TTG GAG CCT ACT ACC TTT CAG GAA CCC CCA Tyr Arg Pro Pro Lys Ile Leu Glu Pro Thr Thr Phe Gln Glu Pro Pro		700		705		710	2343
25	CCC AAG CCC AGC CGG CCC AAG TAC AGA CCT CCT CCA CAG ACC AAC CTG Pro Lys Pro Ser Arg Pro Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu		715		720		725	2391
30	CTG GCT CCT AAG CTG CAG TTC CAG GTC CCT GAG GGT CTG TGT GCC AGC Leu Ala Pro Lys Leu Gln Phe Gln Val Pro Glu Gly Leu Cys Ala Ser		735		740		745	2439
35	TCT CCT ACG CTT ACC AGC CCT ATG GAG TAT CCA TCT CCA GTT AAC TCG Ser Pro Thr Leu Thr Ser Pro Met Glu Tyr Pro Ser Pro Val Asn Ser		750		755		760	2487
40	CTG CAC ACC CCA CCT CTC CAC CGG CAC AAT GTC TTC AAG CGC CAC AGC Leu His Thr Pro Pro Leu His Arg His Asn Val Phe Lys Arg His Ser		765		770		775	2535
45	ATG CGG GAG GAG GAC TTC ATC CGG CCC AGT AGC CGA GAA GAG GCC CAG Met Arg Glu Glu Asp Phe Ile Arg Pro Ser Ser Arg Glu Glu Ala Gln		780		785		790	2583
50	CAG CTC TGG GAG GCA GAG AAG ATC AAG ATG AAG CAG GTC CTA GAA AGA Gln Leu Trp Glu Ala Glu Lys Ile Lys Met Lys Gln Val Leu Glu Arg		795		800		805	2631
55	CAG CAG AAG CAG ATG GTG GAA GAT TCC CAG TGG CTG AGG CGA GAG GAA Gln Gln Lys Gln Met Val Glu Asp Ser Gln Trp Leu Arg Arg Glu Glu		815		820		825	2679
60	AGA TGC TTG GAC CCT ATG GTT TAT ATG AAT GAC AAG TCC CCA CTG ACT Arg Cys Leu Asp Pro Met Val Tyr Met Asn Asp Lys Ser Pro Leu Thr		830		835		840	2727
65	CCA GAG AAG GAG GCC GGC TAC ACG GAG TTC ACA GGG CCC CCA CAG AAA Pro Glu Lys Glu Ala Gly Tyr Thr Glu Phe Thr Gly Pro Pro Gln Lys		845		850		855	2775
70	CCA CCT CGG CTC GGT GCA CAG TCC ATT CAG CCC ACA GCC AAC CTG GAC Pro Pro Arg Leu Gly Ala Gln Ser Ile Gln Pro Thr Ala Asn Leu Asp							2823

	860	865	870	
5	AGG ACC GAT GAC CTC GTG TAC CAC AAT GTC ATG ACC CTG GTG GAG GCT Arg Thr Asp Asp Leu Val Tyr His Asn Val Met Thr Leu Val Glu Ala 875 880 885 890	2871		
10	GTG CTG GGA CTC AAG AAC AAG CTT GGC CAG TTG CCC CCT GAG GAC TAT Val Leu Gly Leu Lys Asn Lys Leu Gly Gln Leu Pro Pro Glu Asp Tyr 895 900 905	2919		
15	GTG GTG GTG GTG AAG AAC GTG GGG CTG AAC CTG CGG AAG CTC ATC GGC Val Val Val Val Lys Asn Val Gly Leu Asn Leu Arg Lys Leu Ile Gly 910 915 920	2967		
20	AGT GTG GAC GAT CTC TTG CCC TCC TTG CCG GCA TCT TCG AGG ACA GAG Ser Val Asp Asp Leu Leu Pro Ser Leu Pro Ala Ser Ser Arg Thr Glu 925 930 935	3015		
25	ATT GAA GGG ACC CAG AAA CTG CTC AAC AAA GAC CTG GCA GAG CTC ATC Ile Glu Gly Thr Gln Lys Leu Leu Asn Lys Asp Leu Ala Glu Leu Ile 940 945 950	3063		
30	AAC AAG ATG AAG TTG GCT CAG CAG AAC GCC GTG ACG TCC CTG AGT GAG Asn Lys Met Lys Leu Ala Gln Gln Asn Ala Val Thr Ser Leu Ser Glu 955 960 965 970	3111		
35	GAC TGC AAG CGG CAG ATG CTC ACA GCG TCC CAT ACC CTG GCT GTG GAT Asp Cys Lys Arg Gln Met Leu Thr Ala Ser His Thr Leu Ala Val Asp 975 980 985	3159		
40	GCC AAG AAC CTG CTG GAT GCT GTG GAC CAA GCC AAG GTT GTG GCT AAT Ala Lys Asn Leu Leu Asp Ala Val Asp Gln Ala Lys Val Val Ala Asn 990 995 1000	3207		
45	CTG GCC CAC CCG CCT GCA GAG T GATCAAGAGA GGGGCCACCT GCCTGCATCT Leu Ala His Pro Pro Ala Glu 1005	3259		
50	TCTGCCCCCA CCTGTCTTGG CATACTTTC CTGCCTTGCC TTTGGTTATT GGTCTTCCAG GGAAAGCTGA GAAGAGTCCA TCCCCCTTGC CACTTTGCAC GACGCCCCCT CTCCCCCAA CCCATCCCAG ACTGTGCTAC TCAGGCTGCA TCTGGACAGA AAGGACTCTG GGCACAGACA CGGGGTGGGG TGACATAGTT CATAGGGGTA CTA CTGCCCAG CCACTCCCTC TTACCCAGC CTGGGTTGCT GGAGCATCAT TGGGGTCATG AGTGTACCCC TAACGGCCAA GATGGCTTTC TGCATGGACA TTTGAGAGCC AGTATTCTC CTTCCTCTTC AGCCCTCAGG GACCCCTGAT ACAGAGGGGA CAGAGAGGGG TTTTATTTGT AGAAAAGCTG TGACATGAGG GCTGGACCTG GCTCTCTTGT ACAGTGTACA TTGGAATTTA TTTAATGTGA GTTTGACCTG GATGGACAGC CAAGGGCCAT AGTCCAGGAG CAAACCAATC CAGTCACAGG ACTCTGTGTT TTTATGGAAC	3319 3379 3439 3499 3559 3619 3679 3739 3799		

TGAGTGCCAC AGGGAAGAAG CAGAGAGTCG GAGGTCAGAA TGGGACTTTG TGCCCTTCCT 3859  
 GCGTTTCTCT TCTCCCTCTT TCCTCTCCCC TCTTTTCTTA CGTCTCCTTT TTCTCCTCCC 3919  
 5 CCTTTTCACA TCTGCTCCCC TCCTCTCTCA TGTCTGTGGA GAACATTTAC CTTCTTCTT 3979  
 TTTGATCGGT GGTGAATTA AAATTATTAC CATTTGCTTT GTGAAAAAAA 4029

10 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1009 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

20 Met Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Val Gly Thr Leu  
 1 5 10 15  
 25 Arg Arg Pro Glu Gly Pro Pro Glu Pro Met Val Val Val Pro Val Asp  
 20 25 30  
 Val Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn  
 35 40 45  
 30 Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln  
 50 55 60  
 Thr Glu Ile Gln Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile  
 65 70 75 80  
 35 Gly Pro Asn Ile Gln Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His  
 85 90 95  
 40 Met Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly  
 100 105 110  
 Glu Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg  
 115 120 125  
 45 Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu  
 130 135 140  
 Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn  
 145 150 155 160  
 50 Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val Ser Glu Gly Met Ala Leu  
 165 170 175  
 55 Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe Phe Lys Asp Met Pro His  
 180 185 190

- 146 -

Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu Leu Leu Glu Lys Glu Val  
 195 200 205  
 5 Gly Leu Asp Leu Phe Phe Pro Lys Gln Met Gln Glu Asn Leu Lys Pro  
 210 215 220  
 Lys Gln Phe Arg Lys Met Ile Gln Gln Thr Phe Gln Gln Tyr Ala Ser  
 225 230 235 240  
 10 Leu Arg Glu Glu Glu Cys Val Met Lys Phe Phe Asn Thr Leu Ala Gly  
 245 250 255  
 Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg Cys Glu Leu Ile Gln Gly  
 260 265 270  
 15 Trp Asn Ile Thr Val Asp Leu Val Ile Gly Pro Lys Gly Ile Arg Gln  
 275 280 285  
 Leu Thr Ser Gln Asp Thr Lys Pro Thr Cys Leu Ala Glu Phe Lys Gln  
 290 295 300  
 Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu Glu Thr Gln Ala Val Leu  
 305 310 315 320  
 25 Gln Leu Gly Ile Glu Gly Ala Pro Gln Ser Leu Ser Ile Lys Thr Ser  
 325 330 335  
 Ser Leu Ala Glu Ala Glu Asn Met Ala Asp Leu Ile Asp Gly Tyr Cys  
 340 345 350  
 30 Arg Leu Gln Gly Glu His Lys Gly Ser Leu Ile Met His Ala Lys Lys  
 355 360 365  
 Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln Ile Pro Thr Leu Asn Leu  
 370 375 380  
 Glu Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp  
 385 390 395 400  
 40 Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro  
 405 410 415  
 Gln Tyr Gly Val Ala Arg Glu Glu Val Val Leu Asn Arg Ile Leu Gly  
 420 425 430  
 45 Glu Gly Phe Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys  
 435 440 445  
 Gly Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr  
 450 455 460  
 Gln Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn  
 465 470 475 480  
 55 Leu Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu  
 485 490 495



	Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His	
	500	505 510
5	Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Pro Thr Leu Val Leu	
	515	520 525
	Tyr Thr Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn	
	530	535 540
10	Cys Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro	
	545	550 555 560
	Glu Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp	
15		565 570 575
	Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met	
	580	585 590
20	Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val	
	595	600 605
	Trp Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln	
25		610 615 620
	Pro Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys	
	625	630 635 640
	Gly Asp Arg Leu Pro Lys Pro Glu Leu Cys Pro Pro Val Leu Tyr Thr	
30		645 650 655
	Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe	
	660	665 670
35	Thr Glu Leu Val Cys Ser Leu Ser Asp Ile Tyr Gln Met Glu Lys Asp	
	675	680 685
	Ile Ala Ile Glu Gln Glu Arg Asn Ala Arg Tyr Arg Pro Pro Lys Ile	
40		690 695 700
	Leu Glu Pro Thr Thr Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro	
	705	710 715 720
	Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln	
45		725 730 735
	Phe Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser	
	740	745 750
50	Pro Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu	
	755	760 765
	His Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe	
	770	775 780
55	Ile Arg Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu	

- 148 -

785                      790                      795                      800  
 Lys Ile Lys Met Lys Gln Val Leu Glu Arg Gln Gln Lys Gln Met Val  
                                  805                      810                      815  
 5    Glu Asp Ser Gln Trp Leu Arg Arg Glu Glu Arg Cys Leu Asp Pro Met  
                                  820                      825                      830  
 10   Val Tyr Met Asn Asp Lys Ser Pro Leu Thr Pro Glu Lys Glu Ala Gly  
                                  835                      840                      845  
 Tyr Thr Glu Phe Thr Gly Pro Pro Gln Lys Pro Pro Arg Leu Gly Ala  
                                  850                      855                      860  
 15   Gln Ser Ile Gln Pro Thr Ala Asn Leu Asp Arg Thr Asp Asp Leu Val  
                                  865                      870                      875                      880  
 Tyr His Asn Val Met Thr Leu Val Glu Ala Val Leu Gly Leu Lys Asn  
                                  885                      890                      895  
 20   Lys Leu Gly Gln Leu Pro Pro Glu Asp Tyr Val Val Val Val Lys Asn  
                                  900                      905                      910  
 25   Val Gly Leu Asn Leu Arg Lys Leu Ile Gly Ser Val Asp Asp Leu Leu  
                                  915                      920                      925  
 Pro Ser Leu Pro Ala Ser Ser Arg Thr Glu Ile Glu Gly Thr Gln Lys  
                                  930                      935                      940  
 30   Leu Leu Asn Lys Asp Leu Ala Glu Leu Ile Asn Lys Met Lys Leu Ala  
                                  945                      950                      955                      960  
 Gln Gln Asn Ala Val Thr Ser Leu Ser Glu Asp Cys Lys Arg Gln Met  
                                  965                      970                      975  
 35   Leu Thr Ala Ser His Thr Leu Ala Val Asp Ala Lys Asn Leu Leu Asp  
                                  980                      985                      990  
 40   Ala Val Asp Gln Ala Lys Val Val Ala Asn Leu Ala His Pro Pro Ala  
                                  995                      1000                      1005  
 Glu

45    (2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids

(B) TYPE: amino acid

50    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

55

## (ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 6

(D) OTHER INFORMATION: /label= OTHER FEATURE

5 /note= "X=F OR Y"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

10 Ser Asp Val Trp Ser Xaa Gly  
1 5

## (2) INFORMATION FOR SEQ ID NO:6:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## 25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

SWRTCNACCC ANSWRWANCC

20

## (2) INFORMATION FOR SEQ ID NO:7:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

35 (ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

45 Asp Leu Ala Ala Arg Asn  
1 5

## (2) INFORMATION FOR SEQ ID NO:8:

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

5 CGACGAYCTN GCNRCNAA 18

(2) INFORMATION FOR SEQ ID NO:9:

10 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Trp Met Ala Pro Glu  
1 5

25 (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
30 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

40 GTACCYTCVG GNGCCATCCA 20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
45 (B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

55 CGGGCCGTGC TGGAGCTCAA 20

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

5

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

15 GTCCGTGAAG ATGACGGCAA

20

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

20

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

25

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

30

AAAGCTGTCA TCGAGATGTC C

21

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

35

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

40

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

45

TCGGTGGGTG CTGGCTGGTA GG

22

## (2) INFORMATION FOR SEQ ID NO:15:

50

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 32 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

55

(ii) MOLECULE TYPE: cDNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATCTGGCACC ACACCTTCTA CAATGAGCTG CG

32

10 (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGTCATACTC CTGCTTGCTG ATCCACATCT GC

32

25 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

30 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGCTTATGGA CTACAAGGAC GACGATGACA GGGG

34

40

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

45 (B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

55 AATTCCTTG TCATCGTCGT CCTTATGGTC CATA

34

What is claimed is:

1. An isolated *RAFTK* nucleic acid molecule from a vertebrate organism.
2. The isolated nucleic acid molecule of claim 1 having a nucleic acid sequence shown in one of SEQ ID NOs:1 or 3, or a complement or fragment thereof.
3. The isolated nucleic acid molecule of claim 2 which has at least 70% homologous to a nucleic acid molecule shown in one of SEQ ID NOs:1 or 3.
4. The isolated nucleic acid molecule of claim 1, which encodes a polypeptide with a *RAFTK* bioactivity.
5. An isolated nucleic acid molecule of claim 1, which encodes a polypeptide shown in SEQ ID NO. 2 or 4.
6. An isolated nucleic acid molecule of claim 1, which is capable of hybridizing to a nucleic acid molecule of one of SEQ ID NOs:1 or 3 under stringent conditions.
7. An isolated nucleic acid molecule of claim 1, which encodes a polypeptide that is at least 70% identical to the polypeptide shown in SEQ ID NOs.: 2 or 4.
8. An isolated nucleic acid molecule of claim 2, which comprises the coding region of one of SEQ ID NOs:1 or 3.
9. An isolated nucleic acid molecule of claim 1, which hybridizes to at least 6 consecutive nucleotides of the *RAFTK* gene shown in one of one of SEQ ID NOs:1 or 3.
10. An isolated nucleic acid molecule of claim 9, which further comprises a label.
11. An expression vector, comprising a nucleic acid molecule of claim 2 operably linked to a transcriptional regulatory sequence.

12. An expression vector of claim 11, which is capable of replicating in a cell.
- 5 13. A host cell transfected with an expression vector of claim 12.
14. A method of making a *RAFTK* polypeptide comprising the steps of:
- a. culturing the cell of claim 13 in an appropriate culture medium to produce a *RAFTK* polypeptide; and
- 10 b. isolating the *RAFTK* polypeptide.
15. A transgenic animal in which expression of a genomic sequence encoding a functional *RAFTK* polypeptide is enhanced, induced, prevented or suppressed.
- 15 16. An isolated polypeptide of a vertebrate organism having a *RAFTK* bioactivity.
17. A polypeptide of claim 16, which is at least 70% homologous to the polypeptide shown in SEQ ID NOs. 2 or 4.
- 20 18. A polypeptide of claim 17, which has a molecular weight of approximately 123kD.
- 25 19. A fusion protein comprising a polypeptide of claim 17 and a second polypeptide, said fusion protein containing a detectable label or a matrix binding domain.
- 30 20. A pharmaceutical preparation comprising a therapeutically effective amount of the polypeptide of claim 17 and a pharmaceutically acceptable carrier.
21. An antibody which is specifically reactive with an epitope of the polypeptide of claim 16.
- 35 22. A method for modulating one or more of growth, differentiation, hematopoiesis, or survival in a cell, comprising treating the cell with an effective amount of an agent which modulates the activity of a *RAFTK* protein thereby altering,



relative to the cell in the absence of the agent, at least one of (i) rate of growth, (ii) differentiation, (iii) hematopoiesis or (iv) survival of the cell.

23. The method of claim 22, wherein the cell is selected from the group  
5 consisting of a mast cell, a melanocyte, and a megakaryocyte.

24. A method for modulating one or more of cell adhesion, migration,  
phagocytosis, or motility of a cell, comprising treating the cell with an effective  
amount of an agent which modulates the activity of a *RAFTK* protein thereby altering,  
10 relative to the cell in the absence of the agent, at least one of (i) cell adhesion, (ii)  
migration, (iii) phagocytosis, or (iv) motility of the cell.

25. A method of claim 24, wherein said cell adhesion is modulated by  
modulating focal adhesion formation in a cell, said method comprising treating the  
15 cell with an effective amount of an agent which modulates the activity of a *RAFTK*  
protein.

26. The method of claim 25, wherein said method is used to treat  
metastasis by a tumor cell.  
20

27. A diagnostic assay for identifying a cell or cells at risk for a disorder  
characterized by unwanted cell proliferation or differentiation, comprising detecting,  
in a cell sample, the presence or absence of a genetic lesion characterized by at least  
one of (i) aberrant modification or mutation of a gene encoding a *RAFTK* protein, and  
25 (ii) mis-expression of said gene; wherein a wild-type form of said gene encodes an  
polypeptide with a *RAFTK* bioactivity.

28. The assay of claim 27, wherein detecting said lesion includes:  
a. providing a diagnostic probe comprising a nucleic acid  
30 including a region of nucleotide sequence which hybridizes to a sense or antisense  
sequence of said gene, or naturally occurring mutants thereof, or 5' or 3' flanking  
sequences naturally associated with said gene;  
b. combining said probe with nucleic acid of said cell sample; and  
c. detecting, by hybridization of said probe to said cellular nucleic  
35 acid, the existence of at least one of a deletion of one or more nucleotides from said  
gene, an addition of one or more nucleotides to said gene, a substitution of one or  
more nucleotides of said gene, a gross chromosomal rearrangement of all or a portion

of said gene, a gross alteration in the level of an mRNA transcript of said gene, or a non-wild type splicing pattern of an mRNA transcript of said gene.

29. A method of preparing differentiated blood cells comprising  
5 modulating the activity of a *RAFTK* protein in a progenitor stem cell.

30. The method of claim 29, wherein the differentiated blood cells are megakaryocytes.

10 31. A population of megakaryocytes prepared using the method of claim 30.

32. A population of platelets prepared using the method of claim 30.

15 33. An assay for screening test compounds to identify compounds which modulate *RAFTK* interaction with cellular proteins, comprising:

a. providing a reaction mixture including a *RAFTK* protein, a *RAFTK*-binding protein, and a test compound; and

20 b. detecting the interaction of the *RAFTK* protein and the *RAFTK*-binding protein.

wherein a statistically significant change in the interaction of the proteins in the presence of the test compound is indicative of the capability of a compound to modulate a bioactivity of a *RAFTK* polypeptide.

25 34. The assay of claim 39, wherein the *RAFTK*-binding protein is selected from the group consisting of paxillin, protein kinase C- $\alpha$ , Protein kinase C- $\delta$ , src, fyn, Grb2, PI3 kinase, and the c-fms receptor, and calpain.

30 35. The assay of claim 33, wherein the detecting step comprises determining the level of phosphorylation of *RAFTK* or the *RAFTK* binding protein.

36. The assay of claim 33, wherein the reaction mixture is selected from the group consisting of a reconstituted protein mixture and a cell lysate.

37. The assay of claim 33, wherein the *RAFTK* protein is a recombinant protein.

38. The assay of claim 33, wherein one or both of the *RAFTK* protein and  
5 *RAFTK*-binding protein is a fusion protein.

39. The assay of claim 33, wherein at least one of the *RAFTK* protein and  
*RAFTK*-binding protein comprises an endogenous detectable label for detecting the  
formation of said complex.

10

40. The method of claim 33, which reaction mixture is a whole cell, and  
interaction of the *RAFTK* protein and *RAFTK*-binding protein is detected in a two  
hybrid assay system.

15 41. A *RAFTK* inhibitor identified using the assay of claim 33.

42. A pharmaceutical preparation comprising (i) the composition of claim  
41 in an amount effective for inhibiting proliferation of a cell, and (ii) a  
pharmaceutically acceptable diluent.

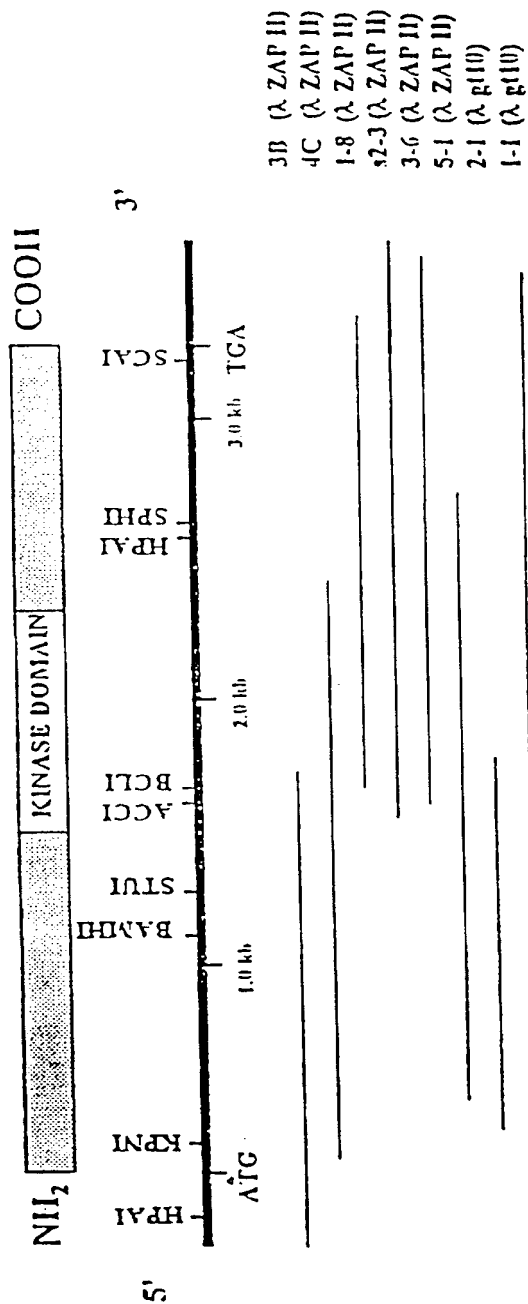
20

43. A method for modulating one or more of growth, differentiation, or  
survival of a megakaryocytic cell, comprising treating the cell with an effective  
amount of the preparation of claim 42 so as to modulate *RAFTK* activity and alter,  
relative to the cell in the absence of the agent, at least one of (i) the growth, (ii)  
25 migration, (iii) differentiation state, or (iv) survival of the cell.

1/5

FIGURE 1

1/5





### FIGURE 3

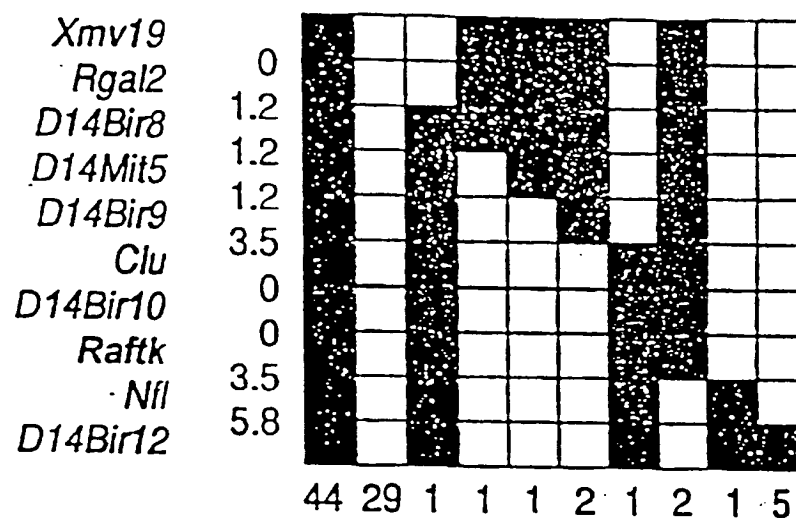
315

RAFTK lak	7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840
--------------	--

## FIGURE 4

B.

Chromosome 14 Rec %



■ B allele inherited from (C57BL/6J x *M. spretus*) F<sub>1</sub> parent  
 □ S allele inherited from (C57BL/6J x *M. spretus*) F<sub>1</sub> parent

5/5

## FIGURE 5

B.

## BXD RI LINE DATA

	<i>Gnrh</i>	<i>Ratfk</i>
01	B	B
02	D	D
05	B	B
06	D	D
08	D	D
09	D	D
11	D	D
12	B	B
13	B	B
14	D	D
15	D	D
16	D	D
18	B	B
19	D	D
20	B	B
21	B	B
22	D	D
23	B	B
24	B	B
25	D	D
27	B	B
28	B	B
29	D	D
30	B	B
31	B	B
32	D	D

*Gnrh*---< 1cM--- *Ratfk*  
95% confidence limits  
0-4.1 Map Units



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/14093

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/52, 15/63, 9/00; C07K 19/00; C12Q 1/68; A61K 38/43;

US CL :536/23.5, 24.31; 435/6, 69.1, 320.1, 325; 530/350; 514/12

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.5, 24.31; 435/6, 69.1, 320.1, 325; 530/350; 514/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN/Medline, HCaPlus; APS/USPAT

search terms: RAFTK, related adhesion focal tyrosine kinase, PYK2, CAK beta

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ---- Y	LEV, S. et al. Protein Tyrosine Kinase PYK2 Involved in Ca <sup>2+</sup> -Induced Regulation of Ion Channel and MAP Kinase Functions. Nature. 31 August 1995. Volume 376, pages 737-745, especially pages 738-740.	1-14, 16-20 ----- 27-28
X ---- Y	AVRAHAM, S. et al. Identification and Characterization of a Novel Related Adhesion Focal Tyrosine Kinase (RAFTK) from Megakaryocytes and Brain. Journal of Biological Chemistry. 17 November 1995. Volume 270, Number 46, pages 27742-27751, especially pages 27742-27747.	1-14, 16-20 ----- 27-28



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*A* document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

14 OCTOBER 1997

Date of mailing of the international search report

06 JAN 1998

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

KAREN E. BROWN

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)\*

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/14093

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SASAKI, H. et al. Cloning and Characterization of Cell Adhesion Kinase Beta, a Novel Protein-Tyrosine Kinase of the Focal Adhesion Kinase Subfamily. Journal of Biological Chemistry. 08 September 1995. Volume 270, Number 36, pages 21206-21219., especially pages 21206-21211.	1-14, 16-20
X	LI, J. et al. Characterization of RAFTK, a Novel Focal Adhesion Kinase, and Its Integrin-Dependent Phosphorylation and Activation in Megakaryocytes. Blood. 15 July 1996. Volume 88, Number 2, pages 417-428, especially pages 418-427.	1-14, 16-20

Form PCT/ISA/210 (continuation of second sheet)(July 1992)★

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/14093

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☒ Claims Nos.: 1-14, 17-20 and 27-28  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
  
Please See Extra Sheet.
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-14, 16-20, 27-28

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/14093

## BOX I. OBSERVATIONS WHERE CLAIMS WERE FOUND UNSEARCHABLE

2. Where no meaningful search could be carried out, specifically:

All of the claims are unsearchable to the extent that they require reference to the specified sequences from the sequence listing. Because the sequence listing supplied by the Applicant was unreadable, no meaningful search of the sequences per se can be carried out by this Authority. However, the subject matter of the claims has been searched to the extent possible with reference to the balance of the description.

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-14, 16-20 and 27-28, drawn to RAFTK polynucleotides, expression vectors, host cells, a method of producing RAFTK recombinantly, RAFTK polypeptide and a method of using RAFTK polynucleotides in a diagnostic assay.

Group II, claim(s) 15, drawn to a transgenic animal in which RAFTK activity is altered.

Group III, claim(s) 21, drawn to an anti-RAFTK antibody.

Group IV, claim(s) 22-26, 29-32 and 43, drawn to a method of altering RAFTK activity using an agent which modulates the effect of RAFTK protein.

Group V, claim(s) 33-40, drawn to a method of using RAFTK protein to identify compounds which modulate RAFTK activity.

Group VI, claim(s) 41-42, drawn to RAFTK inhibitors.

The inventions listed as Groups I-VI do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of Group I is a polynucleotide encoding RAFTK.

Group II does not share the same special technical feature as Group I because the the polynucleotide of Group I is not required for the production of the transgenic animal of Group II.

Group III does not share the same special technical feature as Group I because the polynucleotide of Group I does not encode the antibody of Group III.

Group IV does not share the same special technical feature of Group I because the method of Group IV neither uses nor produces the polynucleotide of Group I, and the polynucleotide of Group I is not required in any steps of the method of Group IV.

Group V does not share the same special technical feature of Group I because the method of Group V neither uses nor produces the polynucleotide of Group I, and the polynucleotide of Group I is not required in any steps of the method of Group V.

Group VI does not share the same special technical feature of Group I because the polynucleotide of Group I is not an inhibitor of Group VI nor does it encode the inhibitor of Group VI, and thus the inhibitor of Group VI does not share the same special technical feature as the polynucleotide of Group I.

Accordingly, the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/52, 15/63, 9/00, C07K 19/00,</b> <b>C12Q 1/68, A61K 38/43</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 98/07870</b> <b>(43) International Publication Date:</b> 26 February 1998 (26.02.98)
<b>(21) International Application Number:</b> PCT/US97/14093 <b>(22) International Filing Date:</b> 12 August 1997 (12.08.97)  <b>(30) Priority Data:</b> 08/703,623 23 August 1996 (23.08.96) US 08/816,462 13 March 1997 (13.03.97) US  <b>(71) Applicant:</b> BETH ISRAEL DEACONESS MEDICAL CENTER, INC. [US/US]; 330 Brookline Avenue, Boston, MA 02215 (US).  <b>(72) Inventors:</b> AVRAHAM, Shalom; 22 Heath Hill, Brookline, MA 02146 (US). AVRAHAM, Hava; 22 Heath Hill, Brookline, MA 02146 (US). GROOPMAN, Jerome, E.; 79 Druce Street, Brookline, MA 02146 (US).  <b>(74) Agents:</b> LOREN, Ralph, A. et al.; Lahive & Cockfield, LLP, 28 State Street, Boston, MA 02109 (US).		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> NOVEL <i>RAFTK</i> SIGNALING MOLECULES AND USES THEREFOR  <b>(57) Abstract</b>  The present invention relates to the discovery of novel " <i>RAFTK</i> " genes and polypeptides. Therapeutics, diagnostics and screening assays based on these molecules are also disclosed.		

\*(Referred to in PCT Gazette No. 21/1998, Section II)

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav	TM	Turkmenistan
BF	Burkina Faso	GR	Greece		Republic of Macedonia	TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's	NZ	New Zealand		
CM	Cameroon		Republic of Korea	PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## NOVEL RAFTK SIGNALING MOLECULES AND USES THEREFOR

### Government Funding

This work was supported, in part, by National Institutes of Health grants: HL 46668, HL 53745-02, HL 55187-01, HL 43510-06A, HL 43510-07, HL 51456-02 and HL 55445-01. The U.S. Government may, therefore, have certain rights to this invention.

### Background of the Invention

Signal transduction is triggered by stimulation of a cell surface receptor which either has kinase activity itself or is physically and/or functionally linked to an intracellular protein tyrosine kinase (PTK) (Cantley, L.C. et al. (1991) *Cell* 64, 281-302; Shattil, S.J., and Brugge, J.S. (1991) *Curr. Opin. Cell Biol.* 3, 869-879; Weiss, A. (1993) *Cell* 73, 209-212). PTKs participate in a variety of signal transduction pathways which modulate cell growth and differentiation (Ullrich, A., and Schlessinger, J. (1990) *Cell* 61, 203-212; Pawson, T., and Gish, G.D. (1992) *Cell* 71, 359-362; Fantl, W.J., et al. (1993) *Ann. Rev. Biochem.* 62, 453-481). Through a series of inducible and reversible protein-protein interactions and phosphorylation-mediated enzymatic activities, protein-tyrosine kinases are recruited to relay signals throughout the cell. Such interactions are involved in all stages of the intracellular signal transduction process - at the plasma membrane, where the signal is initiated; in the cytoplasm, where the signals are disseminated to different cellular locations; and in the nucleus, where other proteins involved in transcriptional control form complexes to regulate transcription of particular genes. Protein kinase cascades allow for amplification, feedback, cross-talk, and branching in signal transduction pathways.

The integrin cell surface receptors are also capable of transducing cytoplasmic signals (Hynes, R.O. (1992) *Cell* 69, 11-25; Juliano, R.L., and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585; Schwartz, M.A. (1992) *Trends Cell Biol.* 2, 304-308) and activation of this pathway is linked to one or more PTKs (Guan, J.-L., et al. (1991) *Cell Regul.* 2, 951-964; Kornberg, L.J. et al. (1991) *Proc. Natl. Acad. Sci. USA* 88, 8392-8396). Integrins are the major family of cell surface receptors that mediate adhesive interactions (Albelda, S. M. and Buck, C. A. (1990) *FASEB J* 4, 2868). Integrin receptor engagement and subsequent clustering of integrins lead to the formation of focal adhesion sites. Protein assemblies of integrins, linked to intracellular cytoskeletal complexes and to bundles of actin filaments in focal adhesions, play critical roles in modulating adhesion and inducing shape changes involved in cell spreading and locomotion (Hynes R O. (1992) *Cell* 69, 11). Such

cellular adhesive interactions, mediated by cell surface receptors that bind to ligands on adjacent cells or in the extracellular matrix, participate in the processes of cell migration, proliferation and differentiation (Gumbiner, B. M. (1993) *Neuron* 11, 551).

5           A large number of cytoplasmic proteins have been identified as components of focal adhesion structures (Burridge K, et al. (1998) *Ann Rev Cell Biol* 4, 487; Turner, C. E. and Burridge, K. (1991) *Curr Opin Cell Biol* 3, 849). These are classified as either structural proteins or signaling molecules. Vinculin,  $\alpha$ -actinin and talin are well known as the major structural proteins at focal adhesion sites (Turner, C. E. and  
10 Burridge, K. (1991) *Curr Opin Cell Biol* 3, 849). *In vitro* and *in vivo* studies have shown that these proteins serve as bridge-like linkages between the integrins and actin filaments, and as a dock for the association of signaling proteins that leads to integrin-induced changes in cell function (Clark, E. A. and Brugge, J.S. (1995) *Science* 268, 233).

15           Several PTKs have been implicated in integrin signaling events by virtue of either their integrin-dependent activation or their localization to these focal contacts (Clark, E. A. and Brugge, J. S. (1995) *Science* 268, 233; Richardson, A. and Parsons, J. T. (1995) *Bioessays* 17, 229). Two focal adhesion proteins that demonstrate a high stoichiometry of tyrosine phosphorylation upon integrin activation are the focal  
20 adhesion kinase (FAK) and paxillin (Schaller, M. D. et al. (1992) *Proc Natl Acad Sci USA* 89, 5192; Hanks, S. K. et al. (1992) *Proc Natl Acad Sci USA* 89, 8487; Burridge, K. et al. (1992) *J Cell Biol* 119, 893). The tyrosine phosphorylation of these two proteins has been suggested as being involved in both the formation of focal adhesions and the assembly of actin stress fibers (Burridge, K. et al. (1992) *J Cell Biol*  
25 119, 893). In addition, the association of FAK with the cytoskeletal protein talin in NIH3T3 cells was observed (Chen, H-C et al. (1995) *J Biol Chem* 270, 16995).

          pp125<sup>FAK</sup> is phosphorylated in response to  $\alpha$ IIb $\beta$ <sub>3</sub>- integrin-mediated cell adhesion (See, e.g., Juliano, R.L., and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585; Hanks, S.K. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 8487-8489; Schaller, M.D., et  
30 al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196). Induction of the kinase activity and the tyrosine phosphorylation of pp125<sup>FAK</sup> were observed following the adherence of fibroblasts to fibronectin (See, e.g., Guan, J.-L. et al. (1991) *Cell Regul.* 2, 951-964; Kornberg, L.J. et al. (1991) *Proc. Natl. Acad. Sci. USA* 88, 8392-8396; Hanks, S.K. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 8487-8489; Schaller, M.D., et  
35 al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196), the adherence of epidermal carcinoma cells to fibronectin, laminin, or collagen type IV (Kornberg, L., et al. (1992) *J. Biol. Chem.* 267, 23439-23442), and the aggregation of platelets in the



presence of fibrinogen, a ligand for  $\alpha\text{IIb}\beta_3$ - integrin (glycoprotein IIb/IIIa) (Lipfert, L., et al. (1992) *J. Cell Biol.* 119, 905-912). Phosphorylated pp125FAK is localized in focal adhesion contacts.

Src tyrosine kinases are capable of interacting with components of focal adhesions upon kinase activation (Weng, Z. et al. (1993) *J Biol Chem* 268, 14956; Schaller, M. D. et al. (1994) *Mol Cell Biol* 14, 1680). The Tyr<sup>397</sup> of FAK is autophosphorylated upon integrin-mediated stimulation. This phosphotyrosine then provides a binding site for the SH2 domain of p60Src and its family members, e.g. p59Fyn (Cobb, B. S. et al. (1994) *Mol Cell Biol* 14, 147 Eide, B. L. et al. (1995) *Mol Cell Biol* 15, 2819). Integrin-mediated signal transduction was found to be linked to the Ras pathway by Grb2 binding to FAK (Schlaepfer, D. D. et al. (1994) *Nature* 372, 786; Kharbanda, S. et al. (1995) *Proc Natl Acad Sci USA* 92, 6132). Phosphatidylinositol 3-kinase (PI-3 kinase) was also associated with the activated FAK (Chen, H-C et al. (1994) *Proc Natl Acad Sci USA* 91, 10148; Guinebault, C. et al. (1995) *J Cell Biol* 129, 831). In addition, FAK phosphorylation is stimulated by a number of other substances, including small peptide mitogens such as vasopressin, bombesin, endothelin (Zachary, I. et al. (1992) *J Biol Chem* 267, 19031; Sinnott-Smith, J. et al. (1993) *J Biol Chem* 268, 14261), and bradykinin (Leeb-Lundberg, L. M. F et al. (1994) *J Biol Chem* 269, 24328); bioactive lipids such as Alzheimer's Ab peptide (Zhang C et al. *J Biol Chem* 269, 25247, 1994); antigens for immunoglobulin E receptors (Hamawy, M. M. et al. (1993) *J Biol Chem* 268, 6851); neuropeptide receptors (Zhang, C. et al. (1994) *J. Biol. Chem.* 269, 25247-25250), growth factors such as hepatocyte growth factor, platelet-derived growth factor and M-CSF-1; (Kharbanda, S. et al. (1995) *Proc Natl Acad Sci USA* 92, 6132; Matsumoto, K. et al. (1994) *J Biol Chem* 269, 31807; Rankin, S. and Rozengurt, E. (1994) *J Biol Chem* 269, 704) and upon oncogenic transformation (Guan, J.-L., and Shalloway, D. (1992) *Nature* 358, 690-692) in adherent cells.

pp125FAK has been cloned from *Xenopus* (*X. Laevis*), avian, rodent, and human species and is expressed in a wide range of cell types (See, e.g., Schaller, M.D. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196; Schaller, M.D., and Parsons, J.T. (1993) *Trends Cell Biol.* 3, 258-262; Clark, E.A., and Brugge, J.S. (1995) *Science* 268, 233-239).

### Summary of the Invention

The present invention is based on the discovery of novel molecules, referred to herein as "related adhesion focal tyrosine kinase" or "RAFTK" polypeptide molecules and the nucleic acid molecules coding therefore. The RAFTK molecules of the

present invention are useful in regulating a variety of cellular processes. The *RAFTK* polypeptide is an intracytoplasmic protein tyrosine kinase.

In one aspect, the invention features isolated vertebrate *RAFTK* nucleic acid molecules. In a preferred embodiment a *RAFTK* nucleic acid has a nucleic acid  
5 sequence shown in one of SEQ ID NOs:1 or 3, or a complement or fragment thereof. The disclosed molecules can be non-coding, (e.g. probe, antisense or ribozyme molecules) or can encode a polypeptide with *RAFTK* bioactivity. In a preferred embodiment a *RAFTK* nucleic acid of the present invention comprises the coding region of one of SEQ ID NOs: 1 or 3. In another preferred embodiment the subject  
10 *RAFTK* nucleic acids encode a polypeptide with a *RAFTK* bioactivity. In a particularly preferred embodiment the nucleic acid of the present invention encodes a polypeptide shown in one of SEQ ID NOs: 2 or 4.

In one embodiment, the nucleic acids of the present invention can hybridize to a vertebrate *RAFTK* gene or to the complement of a vertebrate *RAFTK* gene. In a  
15 further embodiment, a *RAFTK* nucleic acid hybridizes with the coding sequence designated in one of SEQ ID NOs:1 or 3 or to the complement to the coding sequence designated in one of SEQ ID NOs:1 or 3. In a preferred embodiment, the hybridization is conducted under stringent conditions.

In further embodiments, the nucleic acid molecule is a *RAFTK* nucleic acid  
20 molecule that is at least 60%, at least 70%, preferably 80%, more preferably 85%, and even more preferably at least 95% homologous in sequence to the nucleic acids shown in one of SEQ ID NOs:1 or 3 or to the complement of the nucleic acid shown in one of SEQ ID NOs:1 or 3. In another embodiment, the *RAFTK* nucleic acid molecule encodes a polypeptide that is at least 60%, preferably at least 70%, preferably 80%,  
25 and more preferably at least 85%, and even more preferably at least 95% homologous in sequence to the polypeptide shown in one of SEQ ID NOs: 2 or 4.

The invention also provides probes and primers comprising substantially purified oligonucleotides, which correspond to a region of nucleotide sequence which hybridizes to at least 6 consecutive nucleotides of the sequence set forth in one of  
30 SEQ ID NOs:1 or 3, the complement of one of SEQ ID NOs:1 or 3, or naturally occurring mutants thereof. In a preferred embodiment a probe or primer of the present invention hybridizes under stringent conditions to a nucleic acid corresponding to at least 12 consecutive nucleotides of either sense or antisense sequence of one or more of SEQ ID NOs:1 or 3; preferably to at least 25 consecutive nucleotides; and more  
35 preferably to at least 40, 50 or 75 consecutive nucleotides of either sense or antisense sequence of one or more of SEQ ID NOs:1 or 3.

In a preferred embodiment, a probe of the present invention comprises all or a portion of nucleotides 1595-2974 of one of SEQ ID NOs:1 or 3. In preferred embodiments, the probe/primer further includes a label group, which is capable of being detected.

5 For expression, the subject *RAFTK* nucleic acids can include a transcriptional regulatory sequence, e.g. at least one of a transcriptional promoter (e.g., for constitutive expression or inducible expression) or transcriptional enhancer sequence, which regulatory sequence is operably linked to the *RAFTK* gene sequence. Such regulatory sequences in conjunction with a *RAFTK* nucleic acid molecules can be  
10 useful vectors for gene expression. This invention also features host cells transfected with such an expression vector whether prokaryotic or eukaryotic and *in vitro* (e.g. cell culture) and *in vivo* (e.g. transgenic) methods for producing *RAFTK* polypeptides by employing the expression vectors.

The invention also features transgenic non-human animals which include a  
15 heterologous form of a *RAFTK* gene described herein, or which misexpress an endogenous *RAFTK* gene (e.g., an animal in which expression of one or more of the subject *RAFTK* proteins is disrupted). Such a transgenic animal can serve as an animal model for studying cellular and tissue disorders comprising mutated or mis-expressed *RAFTK* alleles or can be used in drug screening. Alternatively, such a  
20 transgenic animal can be useful for expressing recombinant *RAFTK* polypeptides.

In another aspect, the invention features isolated *RAFTK* polypeptides, preferably substantially pure preparations e.g., of plasma purified or recombinantly produced *RAFTK* polypeptides. In preferred embodiments, the polypeptide has a *RAFTK* bioactivity. In addition, *RAFTK* polypeptides which specifically antagonize  
25 the activity of a native *RAFTK* polypeptide, such as can be provided by truncation mutants or other dominant negative mutants, are also specifically contemplated by the present invention.

In one embodiment, the polypeptide is identical to or homologous to a *RAFTK* protein represented in one of SEQ ID NOs: 2 or 4. Related members of the vertebrate  
30 and particularly the mammalian *RAFTK* family are also within the scope of the invention. Preferably, a *RAFTK* polypeptide has an amino acid sequence at least 60%, at least 70% homologous, preferably at least 80%, more preferably at least 90%, and even more preferably at least 95% homologous to the polypeptide represented by one of SEQ ID NOs: 2 or 4. In a preferred embodiment, the *RAFTK* polypeptide is  
35 encoded by a nucleic acid which hybridizes with a nucleic acid sequence represented in one of SEQ ID NOs: 1 or 3. The subject *RAFTK* polypeptides also include modified polypeptides, which are resistant to post-translation modification, as for

example, due to mutations which alter modification sites (such as tyrosine, threonine, serine or asparagine residues), or which prevent glycosylation of the protein, or which prevent interaction of the protein with intracellular proteins.

The *RAFTK* polypeptide can comprise a full length protein, such as  
5 represented in one of SEQ ID NOs: 2 or 4, or it can comprise a fragment corresponding to one or more particular motifs/domains, or to arbitrary sizes, e.g., at least 5, 10, 25, 50, 100, 150 or 200 amino acids in length. In preferred embodiments, the *RAFTK* polypeptide includes at least a portion of an *RAFTK* kinase domain and has a *RAFTK* bioactivity. In preferred embodiments the subject *RAFTK* polypeptides  
10 have a molecular weight of approximately 123kD.

Another aspect of the invention features chimeric molecules (e.g. fusion proteins) comprised of a *RAFTK* polypeptides. For instance, the *RAFTK* polypeptides can be provided as a recombinant fusion protein which includes a second polypeptide portion, e.g., a second polypeptide having an amino acid sequence unrelated  
15 (heterologous) to the *RAFTK* polypeptide, (e.g. the second polypeptide portion is glutathione-S-transferase, an enzymatic activity such as alkaline phosphatase or an epitope tag).

A further aspect of the invention features pharmaceutical preparations including *RAFTK* polypeptides or homologues, or the nucleic acids encoding *RAFTK*  
20 polypeptides and a pharmaceutically acceptable carrier.

Yet another aspect of the present invention pertains to an immunogen comprising a *RAFTK* polypeptide in an immunogenic preparation, the immunogen being capable of eliciting an immune response specific for a *RAFTK* polypeptide, e.g., a humoral response, an antibody response and/or cellular response. In preferred  
25 embodiments, the immunogen comprises an antigenic determinant, e.g. a unique determinant, from the protein represented by one of SEQ ID NOs: 2 or 4.

A still further aspect of the present invention features antibodies and antibody preparations specifically reactive with an epitope of the *RAFTK* protein. In preferred  
30 embodiments the antibody specifically binds to an epitope represented in one of SEQ ID NOs: 2 or 4. In a particularly preferred embodiment, an antibody of the present specifically recognizes amino acids 68-1009 from the *RAFTK* c-terminus.

Yet another aspect of the present invention concerns a method for modulating the growth, migration, differentiation, and/or survival of a cell, e.g., a mast cell, a melanocyte, or a megakaryocyte, by modulating *RAFTK* bioactivity (e.g., by  
35 potentiating or disrupting certain protein-protein interactions in a *RAFTK* signaling pathway). In general, whether carried out *in vivo*, *in vitro*, or *in situ*, the method comprises treating the cell with an effective amount of a *RAFTK* therapeutic so as to

alter, relative to the cell in the absence of treatment, at least one of (i) rate of growth, (ii) differentiation, (iii) hematopoiesis or (iv) survival of the cell. In preferred embodiments the cells are selected from a group including mast cells, melanocytes, and megakaryocytic cells. In another embodiment a *RAFTK* therapeutic can be used in a method of modulating cell adhesion, migration, phagocytosis, or motility. In preferred embodiments, the method can be used to modulate focal adhesion formation or to treat metastasis by a tumor cell.

Accordingly, the method can be carried out with *RAFTK* therapeutics such as peptide and peptidomimetics or other molecules identified in the above-referenced drug screens which agonize or antagonize the effects of signaling from a *RAFTK* protein or ligand binding of a *RAFTK* protein. Other *RAFTK* therapeutics include antisense constructs for inhibiting expression of *RAFTK* proteins, and dominant negative mutants of *RAFTK* proteins which competitively inhibit ligand interactions upstream and signal transduction downstream of the wild-type *RAFTK* protein. In a preferred embodiment the subject *RAFTK* peptides are capable of modulating signal transduction in a pathway involving stem cell factor, thrombin, fibronectin, CSF-1/M-CSF, T cell receptor stimulation, bFGF, oncoprotein M, IL-6, or TNF $\alpha$ .

A further aspect of the present invention provides a method of determining if a subject is at risk for a disorder characterized by unwanted cell proliferation or aberrant control of differentiation. The method includes detecting, in a tissue of the subject, the presence or absence of a genetic lesion characterized by at least one of (i) a mutation of a gene encoding a *RAFTK* protein, e.g. represented in one of SEQ ID NOs: 1 or 3, or a homolog thereof; or (ii) the mis-expression of a *RAFTK* gene. In preferred embodiments, detecting the genetic lesion includes ascertaining the existence of at least one of: a deletion of one or more nucleotides from a *RAFTK* gene; an addition of one or more nucleotides to the gene, a substitution of one or more nucleotides of the gene, a gross chromosomal rearrangement of the gene; an alteration in the level of a messenger RNA transcript of the gene; the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; a non-wild type level of the protein; and/or an aberrant level of *RAFTK* protein.

For example, detecting the genetic lesion can include (i) providing a probe/primer comprised of an oligonucleotide which hybridizes to a sense or antisense sequence of a *RAFTK* gene or naturally occurring mutants thereof, or 5' or 3' flanking sequences naturally associated with the *RAFTK* gene; (ii) contacting the probe/primer with an appropriate nucleic acid containing sample; and (iii) detecting, by hybridization of the probe/primer to the nucleic acid, the presence or absence of the genetic lesion; e.g., wherein detecting the lesion comprises utilizing the probe/primer

to determine the nucleotide sequence of the *RAFTK* gene and, optionally, of the flanking nucleic acid sequences. For example, the primer can be employed in a polymerase chain reaction (PCR) or in a ligation chain reaction (LCR). In alternate embodiments, the level of a *RAFTK* protein is detected in an immunoassay using an antibody which is specifically immunoreactive with the *RAFTK* protein.

In a further aspect the invention provides for methods of preparing differentiated blood cells by modulating the activity of a *RAFTK* protein in a progenitor stem cell. In a preferred embodiment the subject method can be used to prepare megakaryocytes. Platelets can also be prepared using the subject method.

In yet another aspect, the invention provides assays, e.g., for screening test compounds to identify inhibitors, or alternatively, potentiators, of an interaction between a *RAFTK* protein and, for example, an intracellular protein which binds to the *RAFTK* protein. An exemplary method includes the steps of (i) combining a *RAFTK* polypeptide or bioactive fragments thereof, a *RAFTK* target molecule, and a test compound, e.g., under conditions wherein, but for the test compound, the *RAFTK* protein and target molecule are able to interact; and (ii) detecting the formation of a complex which includes the *RAFTK* protein and the target polypeptide either by directly quantitating the complex, or by measuring a bioactivity of the *RAFTK* protein. Several *RAFTK* binding-proteins have been identified and any of these novel interactions can be exploited in the subject drug screening assays. A statistically significant change, such as a decrease, in the interaction of the *RAFTK* and target molecule in the presence of a test compound (relative to what is detected in the absence of the test compound) is indicative of a modulation (e.g., inhibition or potentiation of the interaction between the *RAFTK* protein and the target molecule). In preferred embodiments the ability of a compound to affect the interaction between and one or more of the *RAFTK* binding-proteins selected from the group consisting of paxillin, protein kinase C (PKC)- $\alpha$ , PKC- $\delta$ , src, fyn, Grb2, PI3 kinase, the c-fms receptor, and calpain, is detected. In certain embodiments the phosphorylation state of *RAFTK* or a *RAFTK* binding protein is measured as a readout of protein-protein interaction. In certain embodiments the reaction mixture can be a reconstituted protein mixture or a cell lysate. In certain embodiments the *RAFTK* protein can be a recombinant protein. In certain embodiments either the *RAFTK* protein or the *RAFTK*-binding protein is a fusion protein and in preferred embodiments, at least one of the proteins includes a label group for detection. In yet another embodiment the reaction mixture is a whole cell and the interaction of *RAFTK* and a *RAFTK* binding protein is detected in a two hybrid assay.

In yet another aspect the invention provides for compounds identified using the subject assay, whether agonists or antagonists (inhibitors) of *RAFTK* activity. In one embodiment the compounds identified in the subject screening assays are included in a pharmaceutical preparation. In yet another embodiment the invention provides for a method of modulating cell growth, differentiation or survival by contacting a cell with a pharmaceutical preparation including a compound identified in one of the subject drug screening assays.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

### Brief Description of the Drawings

*Figure 1* is a schematic representation and restriction enzyme map of the *RAFTK* cDNA. The various cDNA clones, obtained from the Human Hippocampus cDNA Library (in Zap II vector) and the CMK/PMA cDNA library (in  $\lambda$ -gt10 vector) are shown as indicated. Restriction enzyme sites are indicated along the length of the cDNA.

*Figure 2* shows a comparison of the deduced amino acid sequence of *RAFTK* with those of m-pp125<sup>FAK</sup>, src, c-fyn, htk and fgfr. Gaps (indicated by dashes) are introduced to optimize the alignment. Amino acid residues found to be conserved are boxed.

*Figure 3* shows an alignment of the predicted amino acid sequences (single-letter code) of the mouse *RAFTK*, human *RAFTK* and the mouse pp125<sup>FAK</sup> gene translated product. Amino acid residues found to be conserved are boxed.

*Figure 4* shows haplotype analysis of Chromosome 14 genetic markers in (C57BL/6J) x M. spretus)F<sub>1</sub> x M. spretus (BSS type) backcross mice showing linkage and relative position of *RAFTK*. Closed boxes indicate the inheritance of the C57BL/6J (B) allele and open boxes indicate the inheritance of the M. spretus (S) allele from the (C57BL/6J) x M. spretus)F<sub>1</sub> parent. Gene names and references to these loci can be found in GBASE. The first two columns indicate the number of backcross progeny with no recombinations. The following columns indicate recombinational events between adjacent loci (signified by a change from open box to closed box). The number of recombinants are listed below each column and the recombination frequency (REC %) between adjacent loci is indicated.

*Figure 5* shows co-segregation of *RAFTK* and Gnrh in BXD RI lines and localization to Chromosome 14. Strain distribution patterns are depicted for *RAFTK* in the BXD RI lines. The RI line distribution pattern is compared with that of the Gnrh

locus. Map units are indicated between *RAFTK* and *Gnrh*, as are 95% confidence limits.

### Detailed Description of the Invention

5 Protein tyrosine kinases (PTKs) play salient roles in a variety of signal transduction pathways which modulate cell growth and differentiation (Ullrich, A., and Schlessinger, J. (1990) *Cell* 61, 203-212; Pawson, T., and Gish, G.D. (1992) *Cell* 71, 359-362; Fantl, W.J., et al. (1993) *Ann. Rev. Biochem.* 62, 453-481). The novel *RAFTK* proteins of the present invention were identified using PCR primers based on  
10 conserved sequences of protein-tyrosine kinases.

The human *RAFTK* gene was cloned from the CMK cell line, which has properties of cells of the megakaryocytic lineage. The 3.6 kb *RAFTK* human cDNA is shown in SEQ ID NO:1. The full length cDNA contains an open reading frame with the first in-frame ATG codon located at nucleotides 294-296, followed by a stop  
15 codon at positions 3260-3262. *RAFTK* is 51% homologous to focal adhesion kinase, FAK at the nucleic acid level.

The murine homolog was subsequently cloned, based on the ability of a probe derived from the human sequence to hybridize to the mouse gene under high stringency conditions. The murine *RAFTK* homolog encodes a protein of 1009 amino  
20 acids. The amino acid sequences of the human and murine *RAFTK* proteins are 95% homologous and the nucleic acid sequences are 90% homologous. The *RAFTK* gene was mapped to human chromosome 8 and to chromosome 14 in the mouse.

The open reading frame of the *RAFTK* nucleic acid encodes a predicted protein of 1009 amino acid residues. The *RAFTK* protein migrates with a molecular  
25 weight of approximately 115-125 kD. In preferred embodiments, the *RAFTK* protein of the present invention is approximately about 123 kD. The *RAFTK* proteins of the present invention can be activated by phosphorylation, and it will be understood that other post-translational modifications can alter the apparent molecular weight of the protein.

30 The *RAFTK* polypeptide contains several structural motifs common to all protein kinases, including the putative ATP-binding site (G<sup>432</sup>-X-G<sup>434</sup>-X-X-G<sup>437</sup>, where X is any amino acid) and three residues that are predicted to interact with the  $\gamma$  phosphate group of the bound ATP molecule (at amino acids 402, 529, and 655). The amino acid sequences at positions 549-554 and 588-592 are also conserved among  
35 protein-tyrosine kinases. The kinase domain consists primarily of the catalytic domain including the putative ATP-binding site (amino acids 432-437). Amino acid residues 880-887 are important in mediating association of *RAFTK* with src and/or



fyn, particularly when the tyrosine residue at 881 is phosphorylated. The kinase domain of *RAFTK* is flanked by N-terminal and C-terminal regions; the N-terminal region of the protein (amino acids 1-39) is unique and the C-terminal region contains a proline-rich stretch (residues 690-767) in which the proline content exceeds 20%.

5 *RAFTK* lacks myristilation sites and SH2 and SH3 domains.

*RAFTK* was found to be expressed in fetal brain, lung, and liver, and to have a less restricted pattern of expression in adults. Expression was detected in a variety of adult tissues, including: CD34+ bone marrow cells, megakaryocytes, platelets, brain (particularly in the amygdala and hippocampus), macrophages, peripheral blood  
10 lymphocytes, spleen, thymus, B lymphocytes, T lymphocytes, and certain cancer cells.

Accordingly, certain aspects of the present invention relate to nucleic acid molecules encoding a vertebrate, e.g., mammalian *RAFTK* polypeptides, the *RAFTK* polypeptides, antibodies immunoreactive with *RAFTK* polypeptides, and preparations  
15 of such compositions. In addition, drug discovery assays are provided for identifying agents which can modulate the biological function of *RAFTK* proteins. Such agents can be useful therapeutically, therefore, to alter the growth and/or differentiation of a cell. Moreover, the present invention provides diagnostic and therapeutic assays and reagents for detecting and treating disorders involving, for example, aberrant  
20 expression (or loss thereof) of mammalian *RAFTK* genes. Other aspects of the invention are described below or will be apparent to those skilled in the art in light of the present disclosure.

### Definitions

25 For convenience, the meaning of certain terms and phrases employed in the specification, examples, and appended claims are provided below.

The term "binding" as used herein is meant to include detectable interactions between molecules, such as can be detected using, for example, a yeast two hybrid  
30 assay. Also encompassed by this term are enzyme/substrate interactions (e.g. phosphorylation). Interactions between a *RAFTK* protein and a *RAFTK*-binding protein can be constitutive, or induced upon stimulation of a cell.

The term "bioactivity" of a *RAFTK* protein is intended to include effects on growth, differentiation, survival, and motility, e.g., migration or adhesion of cells.  
35 *RAFTK* has been shown to have broad involvement in numerous signaling pathways, and to be activated by: stem cell factor, thrombin stimulation, fibronectin, CSF-1/M-CSF, T cell receptor stimulation, bFGF, oncoprotein M, IL-6, and TNFa. *RAFTK* is

also activated by changes intracellular calcium levels, and by activation of protein kinases  $\alpha$  and  $\delta$ . Thus, *RAFTK* is capable of modulating the growth, differentiation, survival, and motility of numerous cell types, including megakaryocytes, T cells, B cells, monocytes, hematopoietic stem cells (e.g., CD34<sup>+</sup> bone marrow cells),  
5 melanocytes, neural cells (particularly in the amygdala and hippocampus), macrophages, peripheral blood lymphocytes, spleen, thymus, B lymphocytes, T lymphocytes, and certain cancer cells (e.g., Kaposi's sarcoma cells). The subject *RAFTK* polypeptides are also capable of modulating platelet function.

*RAFTK* also modulates the formation of focal adhesions and actin stress fibers,  
10 and is thus important in the control of metastatic growth and in the normal cell growth and integrity, and in processes which involve cell motility, such as, for example, phagocytosis. Other bioactivities of the subject *RAFTK* polypeptides are described in more detail herein.

"Cells," "host cells" or "recombinant host cells" are terms used  
15 interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications can occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

20 A "chimeric protein" or "fusion protein" is a fusion of a first amino acid sequence encoding one of the subject mammalian *RAFTK* polypeptides with a second amino acid sequence defining a domain (e.g. polypeptide portion) foreign to and not substantially homologous with any domain of one of the mammalian *RAFTK* polypeptides. A chimeric protein may present a foreign domain which is found (albeit  
25 in a different protein) in an organism which also expresses the first protein, or it may be an "interspecies", "intergenic", etc. fusion of protein structures expressed by different kinds of organisms. In general, a fusion protein can be represented by the general formula X-*RAFTK*-Y, wherein *RAFTK* represents a portion of the protein which is derived from one of the mammalian *RAFTK* proteins, and X and Y are  
30 independently absent or represent amino acid sequences which are not related to one of the mammalian *RAFTK* sequences in an organism, including naturally occurring mutants.

"Complementary" sequences as used herein refer to sequences which have sufficient complementarity to be able to hybridize, forming a stable duplex.

35 A "delivery complex" as used herein refers to a targeting means (e.g. a molecule that results in higher affinity binding of a gene, protein, polypeptide or peptide to a target cell surface and/or increased cellular uptake by a target cell).

Examples of targeting means include: sterols (e.g. cholesterol), lipids (e.g. a cationic lipid, virosome or liposome), viruses (e.g. adenovirus, adeno-associated virus, and retrovirus) or target cell specific binding agents (e.g. ligands recognized by target cell specific receptors).

5 As is well known, genes for a particular polypeptide may exist in single or multiple copies within the genome of an individual. Such duplicate genes may be identical or may have certain modifications, including nucleotide substitutions, additions or deletions, which all still code for polypeptides having substantially the same activity. The term "DNA sequence encoding a mammalian *RAFTK*

10 polypeptide" can refer to one or more genes within a particular individual. Moreover, certain differences in nucleotide sequences can exist between individual organisms, which are called alleles. Such allelic differences can result in differences in amino acid sequence of the encoded polypeptide yet still encode a protein with the same biological activity.

15 As used herein, the term "gene" or "recombinant gene" refers to a nucleic acid molecule comprising an open reading frame encoding one of the mammalian *RAFTK* polypeptides of the present invention, including both exon and (optionally) intron sequences. A "recombinant gene" refers to nucleic acid encoding a mammalian *RAFTK* polypeptide and comprising mammalian *RAFTK*-encoding exon sequences,

20 though it may optionally include intron sequences which are either derived from a chromosomal mammalian *RAFTK* gene or from an unrelated chromosomal gene. Exemplary recombinant genes encoding the subject mammalian *RAFTK* polypeptides are represented in the appended Sequence Listing. The term "intron" refers to a DNA sequence present in a given mammalian *RAFTK* gene which is not translated into

25 protein and is generally found between exons.

"Homology" or "identity" refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then

30 the molecules are homologous or identical at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. An "unrelated" or "non-homologous" sequence shares less than 40% identity, though preferably less than 25% identity, with one of the mammalian *RAFTK* sequences of the present invention.

35 The term "isolated" as used herein with respect to nucleic acids, such as DNA or RNA, refers to molecules separated from other DNAs, or RNAs, respectively, that are present in the natural source of the nucleic acids. For example, an isolated nucleic

acid encoding one of the subject mammalian *RAFTK* polypeptides preferably includes no more than 10 kilobases (kb) of nucleic acid sequence which naturally immediately flanks the mammalian *RAFTK* gene in genomic DNA, more preferably no more than 5kb of such naturally occurring flanking sequences, and most preferably less than 1.5kb of such naturally occurring flanking sequence. Moreover, an "isolated" nucleic acid" includes nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state. The term "isolated" as used herein also refers to a polypeptide that is substantially free of cellular material, viral material, or culture medium when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. Moreover, an "isolated nucleic acid" is meant to include nucleic acid fragments which are not naturally occurring as fragments and would not be found in the natural state. The term "isolated" is also used herein to refer to polypeptides which are isolated from other cellular proteins and is meant to encompass both purified and recombinant polypeptides.

The term "modulation" as used herein refers to both upregulation, i.e., stimulation or potentiation, and downregulation, i.e. suppression, of a response.

The "non-human animals" of the invention include mammals such as rodents, non-human primates, sheep, dog, cow, chickens, amphibians, reptiles, etc. Preferred non-human animals are selected from the rodent family including rat and mouse, most preferably mouse, though transgenic amphibians, such as members of the *Xenopus* genus, and transgenic chickens are also contemplated herein. The term "chimeric animal" is used herein to refer to animals in which the recombinant gene is found, or in which the recombinant is expressed in some but not all cells of the animal. The term "tissue-specific chimeric animal" indicates that one of the recombinant mammalian *RAFTK* genes is present and/or expressed or disrupted in some tissues but not others.

As used herein, the term "nucleic acid" refers to polynucleotides such as deoxyribonucleic acid (DNA), and, where appropriate, ribonucleic acid (RNA). The term should also be understood to include, as equivalents, analogs of either RNA or DNA made from nucleotide analogs, and, as applicable to the embodiment being described, single (sense or antisense) and double-stranded polynucleotides.

The terms "protein", "polypeptide" and "peptide" are used interchangeably herein.

The term "recombinant protein" refers to a polypeptide of the present invention which is produced by recombinant DNA techniques, wherein generally, DNA encoding a mammalian *RAFTK* polypeptide is inserted into a suitable

expression vector which is in turn used to transform a host cell to produce the heterologous protein. Moreover, the phrase "derived from", with respect to a recombinant *RAFTK* gene, is meant to include within the meaning of "recombinant protein" those proteins having an amino acid sequence of a native *RAFTK* protein, or  
5 an amino acid sequence similar thereto which is generated by mutations including substitutions and deletions (including truncation) of a naturally occurring form of the protein.

As used herein, the term "specifically hybridizes" or "specifically detects" refers to the ability of a nucleic acid molecule of the invention to hybridize to at least  
10 approximately 12, 20, 30, 50, 100, 150, 200, or 300 consecutive nucleotides of a vertebrate, preferably mammalian, *RAFTK* gene, such as a *RAFTK* sequence designated in one of SEQ ID NOs:1 or 3, or a sequence complementary thereto, or naturally occurring mutants thereof, such that it shows more than 10 times more hybridization, preferably more than 100 times more hybridization, and even more  
15 preferably more than 100 times more hybridization than it does to to a cellular nucleic acid (e.g., mRNA or genomic DNA) encoding a protein other than a vertebrate, preferably mammalian, *RAFTK* protein as defined herein.

As used herein, the term "tissue-specific promoter" means a DNA sequence that serves as a promoter, i.e., regulates expression of a selected DNA sequence  
20 operably linked to the promoter, and which effects expression of the selected DNA sequence in specific cells of a tissue, such as cells of hepatic or pancreatic origin, neuronal cells, or immune cells. The term also covers so-called "leaky" promoters, which regulate expression of a selected DNA primarily in one tissue, but cause expression in other tissues as well.

"Transcriptional regulatory sequence" is a generic term used throughout the specification to refer to DNA sequences, such as initiation signals, enhancers, and promoters, which induce or control transcription of protein coding sequences with which they are operably linked. In preferred embodiments, transcription of one of the recombinant mammalian *RAFTK* genes is under the control of a promoter sequence  
30 (or other transcriptional regulatory sequence) which controls the expression of the recombinant gene in a cell-type in which expression is intended. It will also be understood that the recombinant gene can be under the control of transcriptional regulatory sequences which are the same or which are different from those sequences which control transcription of the naturally-occurring forms of *RAFTK* proteins.

As used herein, the term "transfection" means the introduction of a nucleic acid, e.g., an expression vector, into a recipient cell by nucleic acid-mediated gene transfer. "Transformation", as used herein, refers to a process in which a cell's

genotype is changed as a result of the cellular uptake of exogenous DNA or RNA, and, for example, the transformed cell expresses a recombinant form of a mammalian *RAFTK* polypeptide or, in the case of anti-sense expression from the transferred gene, the expression of a naturally-occurring form of the *RAFTK* protein is disrupted.

5 As used herein, the term "transgene" means a nucleic acid sequence (encoding, e.g., one of the mammalian *RAFTK* polypeptides, or an transcript which is antisense to a *RAFTK* nucleic acid sequence), which is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but  
10 which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout). A transgene can include one or more transcriptional regulatory sequences and any other nucleic acid, such as introns, that may be necessary for optimal  
15 expression of a selected nucleic acid.

A "transgenic animal" refers to any animal, preferably a non-human mammal, bird or an amphibian, in which one or more of the cells of the animal contain heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the  
20 cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or *in vitro* fertilization, but rather is directed to the introduction of a recombinant DNA molecule. This molecule can be integrated within a chromosome,  
25 or it can be extrachromosomally replicating DNA. In the typical transgenic animals described herein, the transgene causes cells to express a recombinant form of one of the mammalian *RAFTK* proteins, e.g. either agonistic or antagonistic forms. However, transgenic animals in which the recombinant *RAFTK* gene is silent are also contemplated, as for example, the FLP or CRE recombinase dependent constructs  
30 described below. Moreover, "transgenic animal" also includes those recombinant animals in which gene disruption of one or more *RAFTK* genes is caused by human intervention, including both recombination and antisense techniques.

As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. A preferred vector is an  
35 episome, i.e., a nucleic acid capable of extra-chromosomal replication. Preferred vectors are those capable of autonomous replication and/or expression of nucleic acids to which they are linked. Vectors capable of directing the expression of genes to which

they are operatively linked are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of "plasmids" which refer generally to circular double stranded DNA loops which, in their vector form are not bound to the chromosome. In the present specification,  
5 "plasmid" and "vector" are used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors which serve equivalent functions and which become known in the art subsequently hereto.

#### 10 *Nucleic Acids of the Present Invention*

As described below, one aspect of the invention pertains to isolated nucleic acids comprising nucleotide sequences encoding *RAFTK* polypeptides, and/or equivalents of such nucleic acids. The term "equivalent" includes nucleotide sequences encoding functionally equivalent *RAFTK* polypeptides or functionally  
15 equivalent peptides having a bioactivity of a vertebrate *RAFTK* protein such as described herein. Equivalent nucleotide sequences include sequences that differ by one or more nucleotide substitutions, additions or deletions, such as allelic variants; and, therefore, include sequences that differ from the nucleotide sequence of the *RAFTK* gene shown in one of SEQ ID NOs:1 or 3 due to the degeneracy of the  
20 genetic code.

Preferred nucleic acids are vertebrate *RAFTK* nucleic acids. Particularly preferred vertebrate *RAFTK* nucleic acids are mammalian. Regardless of species, *RAFTK* nucleic acids encode polypeptides that are at least 60% similar to an amino acid sequence of a vertebrate *RAFTK*. Preferred nucleic acids encode a *RAFTK*  
25 polypeptide comprising an amino acid sequence at least 60%, at least 70% homologous, preferably at least 80% homologous, more preferably at least 90% homologous with an amino acid sequence of a vertebrate *RAFTK*, e.g., such as a sequence shown in one of SEQ ID NOs:2 or 4. Nucleic acids which encode polypeptides at least about 95%, and even more preferably at least about 98-99%  
30 similarity with an amino acid sequence represented in one of SEQ ID NOs:2 or 4 are most preferred. Still other preferred nucleic acids of the present invention encode a *RAFTK* polypeptide which includes a polypeptide sequence corresponding to all or a portion of amino acid residues of one of SEQ ID NOs:2 or 4, e.g., at least 5, 10, 25, 50, 100, 150 or 200 amino acid residues of that region.

35 Another aspect of the invention provides a nucleic acid which hybridizes under stringent conditions to a nucleic acid represented by one of SEQ ID NOs:1 or 3. Appropriate stringency conditions which promote DNA hybridization, for example,

6.0 x sodium chloride/sodium citrate (SSC) at about 45°C, followed by a wash of 2.0 x SSC at 50°C, are known to those skilled in the art or can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. For example, the salt concentration in the wash step can be selected from a low stringency of about 2.0 x SSC at 50°C to a high stringency of about 0.2 x SSC at 50°C. In addition, the temperature in the wash step can be increased from low stringency conditions at room temperature, about 22°C, to high stringency conditions at about 65°C. Both temperature and salt may be varied, or temperature of salt concentration may be held constant while the other variable is changed. In a particularly preferred embodiment, a *RAFTK* nucleic acid of the present invention binds to one of SEQ ID NOs: 1 or 3 under stringent conditions.

Preferred nucleic acids have a sequence at least 60%, at least 70% homologous and more preferably 80% and even more preferably at least 85% homologous with an amino acid sequence of a mammalian *RAFTK*, e.g., such as a sequence shown in SEQ ID NOs: 1. Nucleic acids at least 90%, more preferably 95%, and most preferably at least about 98-99% homologous with a nucleic sequence represented in SEQ ID NOs: 1 are of course also within the scope of the invention. In preferred embodiments, the nucleic acid is a mammalian *RAFTK* gene and in particularly preferred embodiments, includes all or a portion of the nucleotide sequence corresponding to the coding region of SEQ ID NOs: 1 or 3.

In preferred embodiments, the nucleic acid is a cDNA encoding a polypeptide having at least one bioactivity of the subject *RAFTK* polypeptide.

Nucleic acids having a sequence that differs from the nucleotide sequences shown in one of SEQ ID NOs: 1 or 3 due to degeneracy in the genetic code are also within the scope of the invention. Such nucleic acids encode functionally equivalent peptides (i.e., a peptide having a biological activity of a mammalian *RAFTK* polypeptide) but differ in sequence from the sequence shown in the sequence listing due to degeneracy in the genetic code. For example, a number of amino acids are designated by more than one triplet. Codons that specify the same amino acid, or synonyms (for example, CAU and CAC each encode histidine) may result in "silent" mutations which do not affect the amino acid sequence of a mammalian *RAFTK* polypeptide. However, it is expected that DNA sequence polymorphisms that do lead to changes in the amino acid sequences of the subject *RAFTK* polypeptides will exist among mammals. One skilled in the art will appreciate that these variations in one or more nucleotides (e.g., up to about 3-5% of the nucleotides) of the nucleic acids encoding polypeptides having an activity of a mammalian *RAFTK* polypeptide can exist among individuals of a given species due to natural allelic variation.



As indicated by the examples set out below, *RAFTK* protein-encoding nucleic acids can be obtained from mRNA present in any of a number of eukaryotic cells. Nucleic acids encoding mammalian *RAFTK* polypeptides of the present invention can also be obtained from genomic DNA from both adults and embryos. For example, a gene encoding a *RAFTK* protein can be cloned from either a cDNA or a genomic library in accordance with protocols described herein, as well as those generally known to persons skilled in the art. Examples of tissues and/or libraries suitable for isolation of the subject nucleic acids include brain, thymus, spleen, among others. A cDNA encoding a *RAFTK* protein can be obtained by isolating total mRNA from a cell, e.g. a vertebrate cell, a mammalian cell, or a human cell, including embryonic cells. Double stranded cDNAs can then be prepared from the total mRNA, and subsequently inserted into a suitable plasmid or bacteriophage vector using any one of a number of known techniques. The gene encoding a mammalian *RAFTK* protein can also be cloned using established polymerase chain reaction techniques in accordance with the nucleotide sequence information provided by the invention. The nucleic acid of the invention can be DNA or RNA. A preferred nucleic acid is a cDNA represented by a sequence shown in one of SEQ ID NOs:1 or 3.

#### *Vectors.*

This invention also provides expression vectors containing a nucleic acid encoding a *RAFTK* polypeptide, operably linked to at least one transcriptional regulatory sequence. "Operably linked" is intended to mean that the nucleotide sequence is linked to a regulatory sequence in a manner which allows expression of the nucleotide sequence. Regulatory sequences are art-recognized and are selected to direct expression of the subject mammalian *RAFTK* proteins. Accordingly, the term "transcriptional regulatory sequence" includes promoters, enhancers and other expression control elements. Such regulatory sequences are described in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). In one embodiment, the expression vector includes a recombinant gene encoding a peptide having an agonistic activity of a subject *RAFTK* polypeptide, or alternatively, encoding a peptide which is an antagonistic form of the *RAFTK* protein. Such expression vectors can be used to transfect cells and thereby produce polypeptides, including fusion proteins, encoded by nucleic acids as described herein. Moreover, the gene constructs of the present invention can also be used as a part of a gene therapy protocol to deliver nucleic acids encoding either an agonistic or antagonistic form of one of the subject mammalian *RAFTK* proteins. Thus, another aspect of the invention features expression vectors for *in vivo* or *in vitro* transfection

and expression of a mammalian *RAFTK* polypeptide in particular cell types so as to reconstitute the function of, or alternatively, abrogate the function of *RAFTK* in a tissue. This is desirable, for example, when the naturally-occurring form of the protein is misexpressed; or to deliver a form of the protein which alters survival of tissue. Expression vectors can also be employed to inhibit neoplastic transformation.

In addition to viral transfer methods, such as those illustrated above, non-viral methods can also be employed to cause expression of a subject *RAFTK* polypeptide in the tissue of an animal. Most nonviral methods of gene transfer rely on normal mechanisms used by mammalian cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral targeting means of the present invention rely on endocytic pathways for the uptake of the subject *RAFTK* polypeptide gene by the targeted cell. Exemplary targeting means of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes.

#### 15 *Probes and Primers*

Moreover, the nucleotide sequences determined from the cloning of *RAFTK* genes from mammalian organisms allow for the generation of probes and primers designed for use in identifying and/or cloning *RAFTK* homologs in other cell types, e.g. from other tissues, as well as *RAFTK* homologs from other mammalian organisms. For instance, the present invention also provides a probe/primer comprising a substantially purified oligonucleotide, which oligonucleotide comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least approximately 12, preferably 25, more preferably 40, 50 or 75 consecutive nucleotides of sense or anti-sense sequence of one of SEQ ID NOs:1 or 3, or naturally occurring mutants thereof. For instance, primers based on the nucleic acid represented in one of SEQ ID NOs:1 or 3 can be used in PCR reactions to clone *RAFTK* homologs.

Likewise, probes based on the subject *RAFTK* sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto and able to be detected, e.g. the label group is selected from amongst radioisotopes, fluorescent compounds, enzymes, and enzyme co-factors.

As discussed in more detail below, such probes can also be used as a part of a diagnostic test kit for identifying cells or tissue which misexpress a *RAFTK* protein, such as by measuring a level of a *RAFTK*-encoding nucleic acid in a sample of cells from a patient; e.g. detecting *RAFTK* mRNA levels or determining whether a genomic *RAFTK* gene has been mutated or deleted. Briefly, nucleotide probes can be generated from the subject *RAFTK* genes which facilitate histological screening of

intact tissue and tissue samples for the presence (or absence) of *RAFTK*-encoding transcripts. Similar to the diagnostic uses of anti-*RAFTK* antibodies, the use of probes directed to *RAFTK* messages, or to genomic *RAFTK* sequences, can be used for both predictive and therapeutic evaluation of allelic mutations which might be manifest in, for example, degenerative, neoplastic or hyperplastic disorders (e.g. unwanted cell growth) or abnormal differentiation of tissue. Used in conjunction with immunoassays as described herein, the oligonucleotide probes can help facilitate the determination of the molecular basis for a disorder which may involve some abnormality associated with expression (or lack thereof) of a *RAFTK* protein. For instance, variation in polypeptide synthesis can be differentiated from a mutation in a coding sequence.

#### *Antisense, Ribozyme and Triplex techniques*

One aspect of the invention relates to the use of the isolated nucleic acid in "antisense" therapy. As used herein, "antisense" therapy refers to administration or *in situ* generation of oligonucleotide molecules or their derivatives which specifically hybridize (e.g. bind) under cellular conditions, with the cellular mRNA and/or genomic DNA encoding one or more of the subject *RAFTK* proteins so as to inhibit expression of that protein, e.g. by inhibiting transcription and/or translation. The binding can be by conventional base pair complementarity, or, for example, in the case of binding to DNA duplexes, through specific interactions in the major groove of the double helix. In general, "antisense" therapy refers to the range of techniques generally employed in the art, and includes any therapy which relies on specific binding to oligonucleotide sequences.

An antisense construct of the present invention can be delivered, for example, as an expression plasmid which, when transcribed in the cell, produces RNA which is complementary to at least a unique portion of the cellular mRNA which encodes a mammalian *RAFTK* protein. Alternatively, the antisense construct is an oligonucleotide probe which is generated *ex vivo* and which, when introduced into the cell causes inhibition of expression by hybridizing with the mRNA and/or genomic sequences of a mammalian *RAFTK* gene. Such oligonucleotide probes are preferably modified oligonucleotides which are resistant to endogenous nucleases, e.g. exonucleases and/or endonucleases, and are therefore stable *in vivo*. Exemplary nucleic acid molecules for use as antisense oligonucleotides are phosphoramidate, phosphothioate and methylphosphonate analogs of DNA (see also U.S. Patents 5,176,996; 5,264,564; and 5,256,775). Additionally, general approaches to constructing oligomers useful in antisense therapy have been reviewed, for example,

by Van der Krol et al. (1988) *Biotechniques* 6, 958-976; and Stein et al. (1988) *Cancer Res* 48, 2659-2668.

Antisense approaches involve the design of oligonucleotides (either DNA or RNA) that are complementary to *RAFTK* mRNA. The antisense oligonucleotides  
5 bind to the *RAFTK* mRNA transcripts and prevent translation. Absolute complementarity, although preferred, is not required. A sequence "complementary" to a portion of an RNA, as referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the  
10 case of double-stranded antisense nucleic acids, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid. Generally, the longer the hybridizing nucleic acid, the more base mismatches with an RNA it can contain and still form a stable duplex (or triplex, as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of  
15 standard procedures to determine the melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of the message, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have recently been shown to be effective at  
20 inhibiting translation of mRNAs as well. (Wagner, R. (1994) *Nature* 372, 333). Therefore, oligonucleotides complementary to either the 5' or 3' untranslated, non-coding regions of a *RAFTK* gene could be used in an antisense approach to inhibit translation of endogenous *RAFTK* mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start  
25 codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5', 3' or coding region of *RAFTK* mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In preferred  
30 embodiments, the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides, or at least 50 nucleotides. Oligodeoxyribonucleotides derived from the translation initiation site, e.g., between the -10 and +10 regions of the *RAFTK* nucleotide sequence of interest, are preferred.

The oligonucleotides can be DNA or RNA or chimeric mixtures or derivatives  
35 or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide

may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al. (1989) *Proc. Natl. Acad. Sci. U.S.A.* 86, 6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci.* 84, 648-652; PCT Publication No. W0 88/09810, published December 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No. W0 89/10134, published April 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al. (1988) *BioTechniques* 6, 958-976) or intercalating agents. (See, e.g., Zon (1988) *Pharm. Res.* 5, 539-549). To this end, the oligonucleotide can be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

The antisense oligonucleotide can comprise at least one modified base moiety which is selected from the group including but not limited to 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xantine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5- oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

The antisense oligonucleotide can also comprise at least one modified sugar moiety selected from the group including but not limited to arabinose, 2-fluoroarabinose, xylulose, and hexose.

In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group consisting of a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide is an  $\alpha$ -anomeric oligonucleotide. An  $\alpha$ -anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gautier et al. (1987) *Nucl. Acids Res.* 15, 6625-6641). The

oligonucleotide is a 2'-O-methylribonucleotide (Inoue et al. (1987) *Nucl. Acids Res.* 15, 6131-6148), or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215, 327-330).

Oligonucleotides of the invention can be synthesized by standard methods  
5 known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides can be synthesized by the method of Stein et al. (1988, *Nucl. Acids Res.* 16, 3209), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al. (1988) *Proc.*  
10 *Natl. Acad. Sci. U.S.A.* 85, 7448-7451), etc.

While antisense nucleotides complementary to the *RAFTK* coding region sequence can be used, those complementary to the transcribed untranslated region are most preferred.

The antisense molecules are delivered to cells which express the *RAFTK in vivo*. A number of methods described herein and known in the art can be used for  
15 delivering the subject nucleic acids into to cells. A preferred approach utilizes a recombinant DNA construct in which the antisense oligonucleotide is placed under the control of a strong pol III or pol II promoter. The use of such a construct to transfect target cells in the patient results in the transcription of sufficient amounts of  
20 single stranded RNAs that forms complementary base pairs with the endogenous *RAFTK* transcripts and thereby prevent translation of the *RAFTK* mRNA. For example, a vector can be introduced *in vivo* such that it is taken up by a cell and directs the transcription of an antisense RNA. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the  
25 desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in mammalian cells. Expression of the sequence encoding the antisense RNA can be by any promoter known in the art to act in mammalian, preferably human cells. Such promoters can be inducible or  
30 constitutive. Such promoters include but are not limited to: the SV40 early promoter region (Bernoist and Chambon (1981) *Nature* 290, 304-310), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al. (1980) *Cell* 22, 787-797), the herpes thymidine kinase promoter (Wagner et al. (1981) *Proc. Natl. Acad. Sci. U.S.A.* 78, 1441-1445), the regulatory sequences of the metallothionein  
35 gene (Brinster et al. (1982) *Nature* 296, 39-42), etc. Any type of plasmid, cosmid, YAC or viral vector can be used to prepare the recombinant DNA construct which can be introduced directly into the tissue site; e.g., the choroid plexus or hypothalamus.

Alternatively, viral vectors can be used which selectively infect the desired tissue; (e.g., for brain, herpesvirus vectors may be used), in which case administration may be accomplished by another route (e.g., systematically).

Ribozyme molecules designed to catalytically cleave *RAFTK* mRNA transcripts can also be used to prevent translation of *RAFTK* mRNA and expression of *RAFTK*. (See, e.g., PCT International Publication WO90/11364, published October 4, 1990; Sarver et al. (1990) *Science* 247, 1222-1225). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy *RAFTK* mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988 *Nature*, 334, 585-591. There are numerous potential hammerhead ribozyme cleavage sites within the nucleotide sequence of human *RAFTK* cDNA. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the *RAFTK* mRNA; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in *Tetrahymena Thermophila* (known as the IVS, or L-19 IVS RNA) and which has been extensively described by Thomas Cech and collaborators (Zaug, et al. (1984) *Science*, 224, 574-578; Zaug and Cech (1986) *Science*, 231, 470-475; Zaug, et al. (1986) *Nature*, 324, 429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech (1986) *Cell*, 47, 207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place. The invention encompasses those Cech-type ribozymes which target eight base-pair active site sequences that are present in *RAFTK*.

As in the antisense approach, the ribozymes can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and are delivered to cells which express the *RAFTK* in vivo e.g., T cells. A preferred method of delivery involves using a DNA construct "encoding" the robozyme under the control of a strong constitutive pol III or pol II promoter, so that transfected cells produce sufficient quantities of the ribozyme to destroy endogenous *RAFTK*

and inhibit translation. Because ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

Endogenous *RAFTK* gene expression can also be reduced by inactivating or "knocking out" the *RAFTK* gene or its promoter using targeted homologous recombination. (E.g., see Smithies et al. (1985) *Nature* 317, 230-234; Thomas & Capecchi (1987) *Cell* 51, 503-512; Thompson et al. (1989) *Cell* 5, 313-321; each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional *RAFTK* (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous *RAFTK* gene (either the coding regions or regulatory regions of the *RAFTK* gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express *RAFTK* *in vivo*. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the *RAFTK* gene. Such approaches are particularly suited in the agricultural field where modifications to ES (embryonic stem) cells can be used to generate animal offspring with an inactive *RAFTK* (e.g., see Thomas & Capecchi (1987) and Thompson (1989), *supra*). However this approach can be adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site *in vivo* using appropriate viral vectors, e.g., herpes virus vectors for delivery to brain tissue; e.g., the hypothalamus and/or choroid plexus.

Alternatively, endogenous *RAFTK* gene expression can be reduced by targeting deoxyribonucleotide sequences complementary to the regulatory region of the *RAFTK* gene (i.e., the *RAFTK* promoter and/or enhancers) to form triple helical structures that prevent transcription of the *RAFTK* gene in target cells in the body. (See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6), 569-84; Helene, C., et al. (1992), *Ann. N.Y. Acad. Sci.* 660, 27-36; and Maher, L. J. (1992) *Bioassays* 14(12), 807-15).

Nucleic acid molecules to be used in triple helix formation for the inhibition of transcription are preferably single stranded and composed of deoxyribonucleotides. The base composition of these oligonucleotides promotes triple helix formation via Hoogsteen base pairing rules, which generally require sizable stretches of either purines or pyrimidines to be present on one strand of a duplex. Nucleotide sequences may be pyrimidine-based, which result in TAT and CGC triplets across the three associated strands of the resulting triple helix. The pyrimidine-rich molecules provide base complementarity to a purine-rich region of a single strand of the duplex in a parallel orientation to that strand. In addition, nucleic acid molecules can be chosen that are purine-rich, for example, containing a stretch of g residues. These molecules will form a triple helix with a DNA duplex that is rich in GC pairs, in which the



majority of the purine residues are located on a single strand of the targeted duplex, resulting in CGC triplets across the three strands in the triplex.

Alternatively, the potential sequences that can be targeted for triple helix formation may be increased by creating a so called "switchback" nucleic acid molecule. Switchback molecules are synthesized in an alternating 5'-3', 3'-5' manner, such that they base pair with first one strand of a duplex and then the other, eliminating the necessity for a sizable stretch of either purines or pyrimidines to be present on one strand of a duplex.

Antisense RNA and DNA, ribozyme, and triple helix molecules of the invention may be prepared by any method known in the art for the synthesis of DNA and RNA molecules. These include techniques for chemically synthesizing oligodeoxyribonucleotides and oligoribonucleotides well known in the art such as for example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesize antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Moreover, various well-known modifications to nucleic acid molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribonucleotides or deoxyribonucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

The antisense constructs of the present invention, by antagonizing the normal biological activity of one of the *RAFTK* proteins, can be used in the manipulation of tissue survival, growth, migration, or differentiation, both *in vivo* and *ex vivo*. Furthermore, the anti-sense techniques (e.g. microinjection of antisense molecules, or transfection with plasmids whose transcripts are anti-sense with regard to a *RAFTK* mRNA or gene sequence) can be used to investigate role of *RAFTK* in developmental events, as well as the normal cellular function of *RAFTK* in adult tissue. Such techniques can be utilized in cell culture, but can also be used in the creation of transgenic animals, as detailed below.

*Polypeptides of the Present Invention*

The present invention also makes available isolated *RAFTK* polypeptides which are isolated from, or otherwise substantially free of other cellular proteins, especially other signal transduction factors and/or transcription factors which may normally be associated with the *RAFTK* polypeptide. The term "substantially free of other cellular proteins" (also referred to herein as "contaminating proteins") or "substantially pure or purified preparations" are defined as encompassing preparations of *RAFTK* polypeptides having less than about 20% (by dry weight) contaminating protein, and preferably having less than about 5% contaminating protein. Functional forms of the subject polypeptides can be prepared, for the first time, as purified preparations by using a cloned gene as described herein. The term "purified" when referring to a polypeptide or nucleic acid means that the polypeptide or nucleic acid is present in the substantial absence of other biological macromolecules, such as other proteins. The term "purified" as used herein preferably means at least 80% by dry weight, more preferably in the range of 95-99% by weight, and most preferably at least 99.8% by weight, of biological macromolecules of the same type present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 5000, can be present). The term "pure" as used herein preferably has the same numerical limits as "purified" immediately above. "Isolated" and "purified" do not encompass either natural materials in their native state or natural materials that have been separated into components (e.g., in an acrylamide gel) but not obtained either as pure (e.g. lacking contaminating proteins, or chromatography reagents such as denaturing agents and polymers, e.g. acrylamide or agarose) substances or solutions. In preferred embodiments, purified *RAFTK* preparations lack any contaminating proteins from the same animal from which *RAFTK* is normally produced, as can be accomplished by recombinant expression of, for example, a human *RAFTK* protein in a non-human cell.

Full length proteins or fragments corresponding to one or more particular motifs and/or domains or to arbitrary sizes, for example, at least 5, 10, 25, 50, 75, 100, 125, 150 amino acids in length are within the scope of the present invention.

For example, isolated *RAFTK* polypeptides can include all or a portion of an amino acid sequences corresponding to a *RAFTK* polypeptide represented in one or more of one of SEQ ID NOs:2 or 4 and 4. Isolated peptidyl portions of *RAFTK* proteins can be obtained by screening peptides recombinantly produced from the corresponding fragment of the nucleic acid encoding such peptides. In addition, fragments can be chemically synthesized using techniques known in the art such as

conventional Merrifield solid phase f-Moc or t-Boc chemistry. For example, a *RAFTK* polypeptide of the present invention can be arbitrarily divided into fragments of desired length with no overlap of the fragments, or preferably divided into overlapping fragments of a desired length. The fragments can be produced  
5 (recombinantly or by chemical synthesis) and tested to identify those peptidyl fragments which can function as either agonists or antagonists of a wild-type (e.g., "authentic") *RAFTK* protein.

Another aspect of the present invention concerns recombinant forms of the *RAFTK* proteins. Recombinant polypeptides preferred by the present invention, in  
10 addition to native *RAFTK* proteins, are at least 60% homologous, preferably at least 70% and more preferably at least 80% homologous and even more preferably at least 90 % homologous with an amino acid sequence represented by one of SEQ ID NOs: 2 or 4. Polypeptides which are at least about 95% homologous with a sequence selected from the group consisting of SEQ ID NOs: 2 and 4 are also within the scope of the  
15 invention. In a preferred embodiment, a *RAFTK* protein of the present invention is a mammalian *RAFTK* protein. In a particularly preferred embodiment a *RAFTK* protein comprises the coding sequence of one of SEQ ID NOs:2 or 4. In particularly preferred embodiments, a *RAFTK* protein has a *RAFTK* bioactivity.

In certain preferred embodiments, the invention features a purified or  
20 recombinant *RAFTK* polypeptide having a molecular weight of approximately 115-125kD. In a preferred embodiment, the subject *RAFTK* polypeptide has a molecular weight of 123 kD. It will be understood that certain post-translational modifications, such as phosphorylation, can increase the apparent molecular weight of the *RAFTK* protein relative to the unmodified polypeptide chain.

The present invention further pertains to recombinant forms of one of the  
25 subject *RAFTK* polypeptides which are encoded by genes derived from a mammalian organism, and which have amino acid sequences evolutionarily related to the *RAFTK* proteins represented in one of SEQ ID NOs:2 or 4. Such recombinant *RAFTK* polypeptides preferably are capable of functioning in one of either role of an agonist  
30 or antagonist of at least one biological activity of a wild-type ("authentic") *RAFTK* protein of the appended sequence listing. The term "evolutionarily related to", with respect to amino acid sequences of mammalian *RAFTK* proteins, refers to both polypeptides having amino acid sequences which have arisen naturally, and also to mutational variants of mammalian *RAFTK* polypeptides which are derived, for  
35 example, by combinatorial mutagenesis. Such evolutionarily derived *RAFTK* polypeptides preferred by the present invention have a *RAFTK* bioactivity and are at least 60% homologous, preferably at least 70% homologous, and more preferably at

least 80% homologous and even more preferably at least 90% homologous with the amino acid sequence shown in one of SEQ ID NOs:2 or 4. Polypeptides at least 95-98% homologous are also within the scope of the invention. In a particularly preferred embodiment, a *RAFTK* protein comprises the amino acid coding sequence of one of SEQ ID NOs:2 or 4.

In general, polypeptides referred to herein as having a bioactivity of a mammalian *RAFTK* protein are defined as polypeptides which include an amino acid sequence corresponding (e.g., identical or homologous) to all or a portion of the amino acid sequences of a mammalian *RAFTK* proteins shown in one of SEQ ID NOs:2 or 4 and which mimic or antagonize all or a portion of the biological/biochemical activities of a naturally occurring *RAFTK* protein. Other biological activities of the subject *RAFTK* proteins are described herein or will be reasonably apparent to those skilled in the art. According to the present invention, a polypeptide has biological activity if it is a specific agonist or antagonist of a naturally-occurring form of a mammalian *RAFTK* protein.

The present invention further pertains to methods of producing the subject *RAFTK* polypeptides. For example, a host cell transfected with a nucleic acid vector directing expression of a nucleotide sequence encoding the subject polypeptides can be cultured under appropriate conditions to allow expression of the peptide to occur. The cells may be harvested, lysed and the protein isolated. A cell culture includes host cells, media and other byproducts. Suitable media for cell culture are well known in the art. The recombinant *RAFTK* polypeptide can be isolated from cell culture medium, host cells, or both, using techniques known in the art for purifying proteins including ion-exchange chromatography, gel filtration chromatography, ultrafiltration, electrophoresis, and immunoaffinity purification with antibodies specific for such peptide. In a preferred embodiment, the recombinant *RAFTK* polypeptide is a fusion protein containing a domain which facilitates its purification, such as GST fusion protein or poly(His) fusion protein.

Moreover, it will be generally appreciated that, under certain circumstances, it is advantageous to provide homologs of one of the subject *RAFTK* polypeptides which function in a limited capacity as one of either a *RAFTK* agonist (mimetic) or a *RAFTK* antagonist, in order to promote or inhibit only a subset of the biological activities of the naturally-occurring form of the protein. Thus, specific biological effects can be elicited by treatment with a homolog of limited function, and with fewer side effects relative to treatment with agonists or antagonists which are directed to all of the biological activities of naturally occurring forms of *RAFTK* proteins.

Homologs of each of the subject *RAFTK* proteins can be generated by mutagenesis, such as by discrete point mutation(s), or by truncation. For example, mutation can give rise to homologs which retain substantially the same, or merely a subset, of the biological activity of the *RAFTK* polypeptide from which it was derived.

5 Alternatively, antagonistic forms of the protein can be generated which are able to inhibit the function of the naturally occurring form of the protein, such as by competitively binding to a downstream or upstream member of the *RAFTK* cascade which includes the *RAFTK* protein. In addition, agonistic forms of the protein may be generated which are constitutively active. Thus, the mammalian *RAFTK* protein and

10 homologs thereof provided by the subject invention may be either positive or negative regulators of *RAFTK* activity.

The recombinant *RAFTK* polypeptides of the present invention also include homologs of the authentic *RAFTK* proteins, such as versions of those protein which are resistant to proteolytic cleavage, as for example, due to mutations which alter

15 ubiquitination or other enzymatic targeting associated with the protein.

*RAFTK* polypeptides may also be chemically modified to create *RAFTK* derivatives by forming covalent or aggregate conjugates with other chemical moieties, such as glycosyl groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives of *RAFTK* proteins can be prepared by linking the chemical moieties to

20 functional groups on amino acid sidechains of the protein or at the N-terminus or at the C-terminus of the polypeptide.

Modification of the structure of the subject mammalian *RAFTK* polypeptides can be for such purposes as enhancing therapeutic or prophylactic efficacy, stability (e.g., ex vivo shelf life and resistance to proteolytic degradation in vivo), or post-

25 translational modifications (e.g., to alter phosphorylation pattern of protein). Such modified peptides, when designed to retain at least one activity of the naturally-occurring form of the protein, or to produce specific antagonists thereof, are considered functional equivalents of the *RAFTK* polypeptides described in more detail herein. Such modified peptides can be produced, for instance, by amino acid

30 substitution, deletion, or addition.

For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid (i.e. isosteric and/or isoelectric mutations) will not have a major effect on the

35 biological activity of the resulting molecule. Conservative replacements are those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids can be divided into four families: (1) acidic =

- aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) nonpolar = alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar = glycine, asparagine, glutamine, cysteine, serine, threonine, tyrosine. In similar fashion, the amino acid repertoire can be grouped as (1) acidic = aspartate, glutamate; (2) basic = lysine, arginine, histidine; (3) aliphatic = glycine, alanine, valine, leucine, isoleucine, serine, threonine, with serine and threonine optionally be grouped separately as aliphatic-hydroxyl; (4) aromatic = phenylalanine, tyrosine, tryptophan; (5) amide = asparagine, glutamine; and (6) sulfur-containing = cysteine and methionine. (see, for example, Biochemistry, 2nd ed., Ed. by L. Stryer, WH Freeman and Co.: 1981). Whether a change in the amino acid sequence of a peptide results in a functional *RAFTK* homolog (e.g. functional in the sense that the resulting polypeptide mimics or antagonizes the wild-type form) can be readily determined by assessing the ability of the variant peptide to produce a response in cells in a fashion similar to the wild-type protein, or competitively inhibit such a response.
- Polypeptides in which more than one replacement has taken place can readily be tested in the same manner.

- This invention further provides a method for generating sets of combinatorial mutants of the subject *RAFTK* proteins as well as truncation mutants, and is especially useful for identifying potential variant sequences (e.g. homologs) that modulate a *RAFTK* bioactivity. The purpose of screening such combinatorial libraries is to generate, for example, novel *RAFTK* homologs which can act as either agonists or antagonist, or alternatively, possess novel activities all together. To illustrate, combinatorially-derived homologs can be generated to have an increased potency relative to a naturally occurring form of the protein.

- Likewise, *RAFTK* homologs can be generated by the present combinatorial approach to selectively inhibit (antagonize) an authentic *RAFTK*. For instance, mutagenesis can provide *RAFTK* homologs which are able to bind other proteins in a *RAFTK* signaling pathway yet prevent propagation of the signal, e.g. the homologs can be dominant negative mutants. Moreover, manipulation of certain domains of *RAFTK* by the present method can provide domains more suitable for use in fusion proteins.

- In one embodiment, the variegated library of *RAFTK* variants is generated by combinatorial mutagenesis at the nucleic acid level, and is encoded by a variegated gene library. For instance, a mixture of synthetic oligonucleotides can be enzymatically ligated into gene sequences such that the degenerate set of potential *RAFTK* sequences are expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g. for phage display) containing the set of *RAFTK* sequences therein.

There are many ways by which such libraries of potential *RAFTK* homologs can be generated from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be carried out in an automatic DNA synthesizer, and the synthetic genes then ligated into an appropriate expression vector. The purpose of a degenerate set of genes is to provide, in one mixture, all of the sequences encoding the desired set of potential *RAFTK* sequences. The synthesis of degenerate oligonucleotides is well known in the art (see for example, Narang, S. A. (1983) *Tetrahedron* 39, 3; Itakura et al. (1981) *Recombinant DNA. Proc 3rd Cleveland Sympos. Macromolecules*, ed. AG Walton, Amsterdam: Elsevier pp 273-289; Itakura et al. (1984) *Annu. Rev. Biochem.* 53, 323. Such techniques have been employed in the directed evolution of other proteins (see, for example, Scott et al. (1990) *Science* 249, 386-390; Roberts et al. (1992) *PNAS* 89, 2429-2433; Devlin et al. (1990) *Science* 249, 404-406).

Likewise, a library of coding sequence fragments can be provided for a *RAFTK* clone in order to generate a variegated population of *RAFTK* fragments for screening and subsequent selection of bioactive fragments. A variety of techniques are known in the art for generating such libraries, including chemical synthesis. In one embodiment, a library of coding sequence fragments can be generated by (i) treating a double stranded PCR fragment of a *RAFTK*-coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule; (ii) denaturing the double stranded DNA; (iii) renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products; (iv) removing single stranded portions from reformed duplexes by treatment with S1 nuclease; and (v) ligating the resulting fragment library into an expression vector. By this exemplary method, an expression library can be derived which codes for N-terminal, C-terminal and internal fragments of various sizes.

A wide range of techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a certain property. Such techniques are generally adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of *RAFTK* homologs. The most widely used techniques for screening large gene libraries typically comprise cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates relatively easy isolation of the vector encoding the gene whose product was detected. Each of the illustrative assays

described below are amenable to high through-put analysis as necessary to screen large numbers of degenerate *RAFTK* sequences created by combinatorial mutagenesis techniques.

In one embodiment, cell based assays can be exploited to analyze the  
5 variegated *RAFTK* library. For instance, the library of expression vectors can be transfected into a cell line ordinarily responsive to a ligand which transduces signals via a pathway involving *RAFTK*, such as, for example, stem cell factor, thrombin, fibronectin, CSF-1/M-CSF, T cell receptor stimulation, bFGF, oncoprotein M, IL-6, or TNF $\alpha$ . The transfected cells are then contacted with a ligand and the effect  
10 of the *RAFTK* mutant can be detected, e.g. on cell viability. Plasmid DNA can then be recovered from the cells which score for inhibition, or alternatively, potentiation of a *RAFTK* activity, and the individual clones further characterized.

Combinatorial mutagenesis has a potential to generate very large libraries of mutant proteins, e.g., in the order of  $10^{26}$  molecules. Combinatorial libraries of this  
15 size may be technically challenging to screen even with high throughput screening assays. To overcome this problem, a new technique has been developed recently, recursive ensemble mutagenesis (REM), which allows one to avoid the very high proportion of non-functional proteins in a random library and simply enhances the frequency of functional proteins, thus decreasing the complexity required to achieve a  
20 useful sampling of sequence space. REM is an algorithm which enhances the frequency of functional mutants in a library when an appropriate selection or screening method is employed (Arkin and Yourvan, 1992, *PNAS USA* 89, 7811-7815; Yourvan et al. (1992) *Parallel Problem Solving from Nature*, 2, In Maenner and Manderick, eds., Elsevir Publishing Co., Amsterdam, pp. 401-410; Delgrave et al.  
25 (1993) *Protein Engineering* 6(3), 327-331).

The invention also provides for reduction of the mammalian *RAFTK* proteins to generate mimetics, e.g. peptide or non-peptide agents, which are able to disrupt  
binding of a mammalian *RAFTK* polypeptide of the present invention with either  
30 upstream or downstream components of a TGF $\beta$  signaling cascade, such as binding proteins or interactors. Thus, such mutagenic techniques as described above are also useful to map the determinants of the *RAFTK* proteins which participate in protein-protein interactions involved in, for example, binding of the subject mammalian *RAFTK* polypeptide to proteins which may function upstream (including both  
35 activators and repressors of its activity) or to proteins or nucleic acids which may function downstream of the *RAFTK* polypeptide, whether they are positively or negatively regulated by it. To illustrate, the critical residues of a subject *RAFTK* polypeptide which are involved in molecular recognition of binding proteins upstream



or downstream of a *RAFTK* can be determined and used to generate *RAFTK*-derived peptidomimetics which competitively inhibit binding of the authentic *RAFTK* protein with that moiety. By employing, for example, scanning mutagenesis to map the amino acid residues of each of the subject *RAFTK* proteins which are involved in binding other extracellular proteins, peptidomimetic compounds can be generated which mimic those residues of the *RAFTK* protein which facilitate the interaction. Such mimetics may then be used to interfere with the normal function of a *RAFTK* protein. For instance, non-hydrolyzable peptide analogs of such residues can be generated using benzodiazepine (e.g., see Freidinger et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), azepine (e.g., see Huffman et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), substituted gamma lactam rings (Garvey et al. in *Peptides: Chemistry and Biology*, G.R. Marshall ed., ESCOM Publisher: Leiden, Netherlands, 1988), keto-methylene pseudopeptides (Ewenson et al. (1986) *J Med Chem* 29, 295; and Ewenson et al. in *Peptides: Structure and Function (Proceedings of the 9th American Peptide Symposium)* Pierce Chemical Co. Rockland, IL, 1985), b-turn dipeptide cores (Nagai et al. (1985) *Tetrahedron Lett* 26, 647; and Sato et al. (1986) *J Chem Soc Perkin Trans* 1, 1231), and b-aminoalcohols (Gordon et al. (1985) *Biochem Biophys Res Commun* 126, 419; and Dann et al. (1986) *Biochem Biophys Res Commun* 134, 71).

#### *Cells expressing recombinant RAFTK polypeptides.*

This invention also pertains to a host cell transfected to express a recombinant form of the subject *RAFTK* polypeptides. The host cell can be any prokaryotic or eukaryotic cell. Thus, a nucleotide sequence derived from the cloning of mammalian *RAFTK* proteins, encoding all or a selected portion of the full-length protein, can be used to produce a recombinant form of a mammalian *RAFTK* polypeptide via microbial or eukaryotic cellular processes. Ligating the polynucleotide sequence into a gene construct, such as an expression vector, and transforming or transfecting into hosts, either eukaryotic (yeast, avian, insect or mammalian) or prokaryotic (bacterial cells), are standard procedures used in producing other well-known proteins, e.g. MAP kinase, p53, WT1, PTP phosphatases, SRC, and the like. Similar procedures, or modifications thereof, can be employed to prepare recombinant *RAFTK* polypeptides by microbial means or tissue-culture technology in accord with the subject invention.

The recombinant *RAFTK* genes can be produced by ligating nucleic acid encoding a *RAFTK* protein, or a portion thereof, into a vector suitable for expression in either prokaryotic cells, eukaryotic cells, or both. Expression vectors for

production of recombinant forms of the subject *RAFTK* polypeptides include plasmids and other vectors. For instance, suitable vectors for the expression of a *RAFTK* polypeptide include plasmids of the types: pBR322-derived plasmids, pEMBL-derived plasmids, pEX-derived plasmids, pBTac-derived plasmids and pUC-derived plasmids for expression in prokaryotic cells, such as *E. coli*.

A number of vectors exist for the expression of recombinant proteins in yeast. For instance, YEP24, YIP5, YEP51, YEP52, pYES2, and YRP17 are cloning and expression vehicles useful in the introduction of genetic constructs into *S. cerevisiae* (see, for example, Broach et al. (1983) in *Experimental Manipulation of Gene Expression*, ed. M. Inouye Academic Press, p. 83, incorporated by reference herein). These vectors can replicate in *E. coli* due the presence of the pBR322 ori, and in *S. cerevisiae* due to the replication determinant of the yeast 2 micron plasmid. In addition, drug resistance markers such as ampicillin can be used. In an illustrative embodiment, a *RAFTK* polypeptide is produced recombinantly utilizing an expression vector generated by sub-cloning the coding sequence of one of the *RAFTK* genes represented in one of SEQ ID NOs:1 or 3.

The preferred mammalian expression vectors contain both prokaryotic sequences, to facilitate the propagation of the vector in bacteria, and one or more eukaryotic transcription units that are expressed in eukaryotic cells. The pcDNA1/amp, pcDNA1/neo, pRc/CMV, pSV2gpt, pSV2neo, pSV2-dhfr, pTk2, pRSVneo, pMSG, pSVT7, pko-neo and pHyg derived vectors are examples of mammalian expression vectors suitable for transfection of eukaryotic cells. Some of these vectors are modified with sequences from bacterial plasmids, such as pBR322, to facilitate replication and drug resistance selection in both prokaryotic and eukaryotic cells. Alternatively, derivatives of viruses such as the bovine papillomavirus (BPV-1), or Epstein-Barr virus (pHEBo, pREP-derived and p205) can be used for transient expression of proteins in eukaryotic cells. The various methods employed in the preparation of the plasmids and transformation of host organisms are well known in the art. For other suitable expression systems for both prokaryotic and eukaryotic cells, as well as general recombinant procedures, see *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989) Chapters 16 and 17.

In some instances, it is desirable to express the recombinant *RAFTK* polypeptide by the use of a baculovirus expression system. Examples of such baculovirus expression systems include pVL-derived vectors (such as pVL1392, pVL1393 and pVL941), pAcUW-derived vectors (such as pAcUW1), and pBlueBac-derived vectors (such as the  $\beta$ -gal containing pBlueBac III).

*Fusion proteins and Immunogens.*

In another embodiment, the coding sequences for the polypeptide can be incorporated as a part of a fusion gene including a nucleotide sequence encoding a different polypeptide. This type of expression system can be useful under conditions where it is desirable to produce an immunogenic fragment of a *RAFTK* protein. For example, the VP6 capsid protein of rotavirus can be used as an immunologic carrier protein for portions of the *RAFTK* polypeptide, either in the monomeric form or in the form of a viral particle. The nucleic acid sequences corresponding to the portion of a subject *RAFTK* protein to which antibodies are to be raised can be incorporated into a fusion gene construct which includes coding sequences for a late vaccinia virus structural protein to produce a set of recombinant viruses expressing fusion proteins comprising *RAFTK* epitopes as part of the virion. It has been demonstrated with the use of immunogenic fusion proteins utilizing the Hepatitis B surface antigen fusion proteins that recombinant Hepatitis B virions can be utilized in this role as well. Similarly, chimeric constructs coding for fusion proteins containing a portion of a *RAFTK* protein and the poliovirus capsid protein can be created to enhance immunogenicity of the set of polypeptide antigens (see, for example, EP Publication No: 0259149; and Evans et al. (1989) *Nature* 339, 385; Huang et al. (1988) *J. Virol.* 62, 3855; and Schlienger et al. (1992) *J. Virol.* 66, 2).

The Multiple Antigen Peptide system for peptide-based immunization can also be utilized to generate an immunogen, wherein a desired portion of a *RAFTK* polypeptide is obtained directly from organo-chemical synthesis of the peptide onto an oligomeric branching lysine core (see, for example, Posnett et al. (1988) *JBC* 263, 1719 and Nardelli et al. (1992) *J. Immunol.* 148, 914). Antigenic determinants of *RAFTK* proteins can also be expressed and presented by bacterial cells.

In addition to utilizing fusion proteins to enhance immunogenicity, it is widely appreciated that fusion proteins can also facilitate the expression of proteins, and accordingly, can be used in the expression of the mammalian *RAFTK* polypeptides of the present invention. For example, *RAFTK* polypeptides can be generated as glutathione-S-transferase (GST-fusion) proteins. Such GST-fusion proteins can enable easy purification of the *RAFTK* polypeptide, as for example by the use of glutathione-derivatized matrices (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. (N.Y.: John Wiley & Sons, 1991)).

In another embodiment, a fusion gene coding for a purification leader sequence, such as a poly-(His)/enterokinase cleavage site sequence at the N-terminus of the desired portion of the recombinant protein, can allow purification of the

expressed fusion protein by affinity chromatography using a Ni<sup>2+</sup> metal resin. The purification leader sequence can then be subsequently removed by treatment with enterokinase to provide the purified protein (e.g., see Hochuli et al. (1987) *J. Chromatography* 411, 177; and Janknecht et al. *PNAS* 88, 8972).

- 5           Techniques for making fusion genes are known to those skilled in the art. Essentially, the joining of various DNA fragments coding for different polypeptide sequences is performed in accordance with conventional techniques, employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide  
10           for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can  
15           be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. John Wiley & Sons: 1992).

#### *Antibodies*

- Another aspect of the invention pertains to antibodies specifically reactive  
20           with a vertebrate *RAFTK* protein, preferably a mammalian *RAFTK* protein. For example, by using immunogens derived from a *RAFTK* protein, e.g. based on the cDNA sequences, anti-protein/anti-peptide antisera or monoclonal antibodies can be made by standard protocols (See, for example, *Antibodies: A Laboratory Manual* ed. by Harlow and Lane (Cold Spring Harbor Press: 1988)). A mammal, such as a  
25           mouse, a hamster or rabbit can be immunized with an immunogenic form of the peptide (e.g., a mammalian *RAFTK* polypeptide or an antigenic fragment which is capable of eliciting an antibody response, or a fusion protein as described above). Techniques for conferring immunogenicity on a protein or peptide include conjugation to carriers or other techniques well known in the art. An immunogenic  
30           portion of a *RAFTK* protein can be administered in the presence of adjuvant. The progress of immunization can be monitored by detection of antibody titers in plasma or serum. Standard ELISA or other immunoassays can be used with the immunogen as antigen to assess the levels of antibodies. In a preferred embodiment, the subject antibodies are immunospecific for antigenic determinants of a *RAFTK* protein of a  
35           mammal, e.g. antigenic determinants of a protein represented by one of SEQ ID NOs:2 or 4.

Following immunization of an animal with an antigenic preparation of a *RAFTK* polypeptide, anti- *RAFTK* antisera can be obtained and, if desired, polyclonal anti- *RAFTK* antibodies isolated from the serum. To produce monoclonal antibodies, antibody-producing cells (lymphocytes) can be harvested from an immunized animal and fused by standard somatic cell fusion procedures with immortalizing cells such as myeloma cells to yield hybridoma cells. Such techniques are well known in the art, and include, for example, the hybridoma technique (originally developed by Kohler and Milstein, (1975) *Nature*, 256, 495-497), the human B cell hybridoma technique (Kozbar et al., (1983) *Immunology Today*, 4, 72), and the EBV-hybridoma technique to produce human monoclonal antibodies (Cole et al., (1985) *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. pp. 77-96). Hybridoma cells can be screened immunochemically for production of antibodies specifically reactive with a mammalian *RAFTK* polypeptide of the present invention and monoclonal antibodies isolated from a culture comprising such hybridoma cells.

The term "antibody" as used herein is intended to include fragments thereof which are also specifically reactive with one of the subject mammalian *RAFTK* polypeptides. Antibodies can be fragmented using conventional techniques and the fragments screened for utility in the same manner as described above for whole antibodies. For example, F(ab)<sub>2</sub> fragments can be generated by treating antibody with pepsin. The resulting F(ab)<sub>2</sub> fragment can be treated to reduce disulfide bridges to produce Fab fragments. The antibody of the present invention is further intended to include bispecific and chimeric molecules having affinity for a *RAFTK* protein conferred by at least one CDR region of the antibody.

Antibodies which specifically bind *RAFTK* epitopes can also be used in immunohistochemical staining of tissue samples in order to evaluate the abundance and pattern of expression of each of the subject *RAFTK* polypeptides. Anti-*RAFTK* antibodies can be used diagnostically in immuno-precipitation and immuno-blotting to detect and evaluate *RAFTK* protein levels in tissue as part of a clinical testing procedure. Likewise, the ability to monitor *RAFTK* protein levels in an individual can allow determination of the efficacy of a given treatment regimen for an individual afflicted with such a disorder. Diagnostic assays using anti- *RAFTK* antibodies can include, for example, immunoassays designed to aid in early diagnosis of a degenerative disorder, particularly ones which are manifest at birth. Diagnostic assays using anti- *RAFTK* polypeptide antibodies can also include immunoassays designed to aid in early diagnosis and phenotyping neoplastic or hyperplastic disorders.

Another application of anti-*RAFTK* antibodies of the present invention is in the immunological screening of cDNA libraries constructed in expression vectors

such as  $\lambda$ gt11,  $\lambda$ gt18-23,  $\lambda$ ZAP, and  $\lambda$ ORF8. Messenger libraries of this type, having coding sequences inserted in the correct reading frame and orientation, can produce fusion proteins. For instance,  $\lambda$ gt11 will produce fusion proteins whose amino termini consist of  $\beta$ -galactosidase amino acid sequences and whose carboxy termini consist of a foreign polypeptide. Antigenic epitopes of a *RAFTK* protein, e.g. other orthologs of a particular *RAFTK* protein or other paralogs from the same species, can then be detected with antibodies, as, for example, reacting nitrocellulose filters lifted from infected plates with anti-*RAFTK* antibodies. Positive phage detected by this assay can then be isolated from the infected plate. Thus, the presence of *RAFTK* homologs can be detected and cloned from other animals, as can alternate isoforms (including splicing variants) from humans.

#### *Methods of Treating Disease*

There are a wide variety of pathological conditions which can be treated using the *RAFTK* therapeutics of the present invention. For example, *RAFTK* therapeutics which modulate *RAFTK* activity in B cells, T cells, and monocytes can be used to treat immune-mediated disorders and mediate both cell mediated and humoral immune responses.

Normal hematopoietic cells are dependent on growth factors for growth and differentiation and the loss of this growth factor dependence can lead to autonomous growth. The involvement of *RAFTK* in several growth factor signaling pathways indicates that missexpression of *RAFTK* can lead to the development of cancers, and the present invention contemplates modulating *RAFTK* expression and/or activity to control aberrant cell growth. In a preferred embodiment *RAFTK* is modulated to treat cancers of hematopoietic cells. In another embodiment malignancy can be suppressed in certain cells e.g., leukemic cells, by modulating *RAFTK* to induce cellular differentiation in "differentiation therapy", for example, in the treatment of leukemia, as has been demonstrated with cytokines or other compounds (Sachs (1996) *Proc. Natl. Acad. Sci. USA* 93:4742).

The subject *RAFTK* proteins can also be modulated to either induce or inhibit apoptosis in a cell. In certain embodiments the subject *RAFTK* proteins can be manipulated to induce apoptosis in cancer cells. In some embodiments *RAFTK* can be modulated in a patient in conjunction with other cancer therapies. Alternatively, in instances when it is desirable to inhibit apoptosis, such as apoptosis induced by chemotherapeutic compounds and irradiation, *RAFTK* may be modulated to inhibit apoptosis.

Cytoskeletal rearrangement has been correlated with growth control and gene expression and is critical in cell adhesion and migration and the modulation of *RAFTK* bioactivity can alter cellular functions which depend upon the cytoskeleton, including, for example, normal tissue maintenance and proliferation and tissue remodeling which occur in response to injury (Turner et al. (1995) *J. Cell Science* 108:333). In a preferred embodiment, *RAFTK* bioactivity is modulated to reduce metastasis of a cancer cell.

Yet another aspect of the present invention pertains to the therapeutic application of a *RAFTK* therapeutic to enhance survival of neurons and other neuronal cells in both the central nervous system and the peripheral nervous system. The expression of *RAFTK* in neuronal cells and their role in signaling pathways involved in apoptosis (Tokiwa et al. (1996) *Science* 273:792) indicates that certain of the *RAFTK* proteins participate in control of adult neurons with regard to maintenance, functional performance, and aging of normal cells; repair and regeneration processes in chemically or mechanically lesioned cells; and prevention of degeneration and premature death which result from loss of differentiation in certain pathological conditions. The present invention specifically contemplates applications of the subject method to the treatment of (prevention and/or reduction of the severity of) neurological conditions deriving from: (i) acute, subacute, or chronic injury to the nervous system, including traumatic injury, chemical injury, vascular injury and deficits (such as the ischemia resulting from stroke), together with infectious/ inflammatory and tumor-induced injury; (ii) aging of the nervous system including Alzheimer's disease; (iii) chronic neurodegenerative diseases of the nervous system, including Parkinson's disease, Huntington's chorea, amyotrophic lateral sclerosis and the like, as well as spinocerebellar degenerations; and (iv) chronic immunological diseases of the nervous system or affecting the nervous system, including multiple sclerosis.

Many neurological disorders are associated with degeneration of discrete populations of neuronal elements and can be treatable with a therapeutic regimen which includes a *RAFTK* therapeutic. For example, Alzheimer's disease is associated with deficits in several neurotransmitter systems, both those that project to the neocortex and those that reside with the cortex. For instance, the nucleus basalis in patients with Alzheimer's disease have been observed to have a profound (75%) loss of neurons compared to age-matched controls. Although Alzheimer's disease is by far the most common form of dementia, several other disorders can produce dementia. Several of these are degenerative diseases characterized by the death of neurons in various parts of the central nervous system, especially the cerebral cortex. However, some forms of dementia are associated with degeneration of the thalamus or the white

matter underlying the cerebral cortex. Here, the cognitive dysfunction results from the isolation of cortical areas by the degeneration of efferents and afferents.

Huntington's disease involves the degeneration of intrastriatal and cortical cholinergic neurons and GABAergic neurons. Pick's disease is a severe neuronal degeneration in the neocortex of the frontal and anterior temporal lobes, sometimes accompanied by death of neurons in the striatum. Treatment of patients suffering from such degenerative conditions can include the application of *RAFTK* therapeutics, in order to control, for example, differentiation and apoptotic events which give rise to loss of neurons (e.g. to enhance survival of existing neurons) as well as promote differentiation and repopulation by progenitor cells in the area affected.

In addition to degenerative-induced dementias, a pharmaceutical preparation of one or more of the subject *RAFTK* therapeutics can be applied opportunely in the treatment of neurodegenerative disorders which have manifestations of tremors and involuntary movements. Parkinson's disease, for example, primarily affects subcortical structures and is characterized by degeneration of the nigrostriatal pathway, raphe nuclei, locus cereleus, and the motor nucleus of vagus. Ballism is typically associated with damage to the subthalamic nucleus, often due to acute vascular accident.

Also included in the methods of the invention are treatment of neurogenic and myopathic diseases which ultimately affect the somatic division of the peripheral nervous system and are manifest as neuromuscular disorders. In an illustrative embodiment, the subject method is used to treat amyotrophic lateral sclerosis. ALS is a name given to a complex of disorders that comprise upper and lower motor neurons. Patients may present with progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, or a combination of these conditions. The major pathological abnormality is characterized by a selective and progressive degeneration of the lower motor neurons in the spinal cord and the upper motor neurons in the cerebral cortex. The therapeutic application of a *RAFTK* therapeutic, can be used alone, or in conjunction with neurotrophic factors such as CNTF, BDNF or NGF to prevent and/or reverse motor neuron degeneration in ALS patients.

Another aspect of the present invention relates to a method of inducing and/or maintaining a differentiated state or enhancing survival by contacting the cells with an agent which modulates *RAFTK*-dependent signaling by a growth factor. For instance, it is contemplated by the invention that, in light of the present finding of a broad involvement of *RAFTK* proteins in signal transduction in a variety of different cell types, the subject *RAFTK* signaling molecules can be used in a wide range of therapeutic regimens both *in vitro* and *in vivo*. A "*RAFTK* therapeutic" can be, as



appropriate, any of the preparations described above, including isolated polypeptides, gene therapy constructs, antisense molecules, peptidomimetics or agents identified in the drug assays provided herein.

5 In one embodiment the *RAFTK* proteins of the present invention can modulate the differentiation or maturation of hematopoietic cells; the subject *RAFTK* polypeptides are capable of influencing both the differentiation and maturation of pluripotent stem cells and the proliferation of differentiated cells. In a preferred embodiment *RAFTK* bioactivity is modulated in CD34+ bone marrow cells; the presence of the cell-surface marker CD34 in humans has been found to correlate with  
10 bone marrow progenitors which proliferate to hematopoietic cytokines.

Hematopoiesis can be modulated either *in vitro* or *in vivo* and the subject *RAFTK* therapeutics can be used alone or in combination with cytokines and/or colony stimulating factors. For example, in certain embodiments, it may be desirable to coadminister a growth factor, for example, G-CSF and/or IL-3 (Lemoli et al. Experimental Hematology 1995, 23:1520) or SCF which has been shown to act with  
15 other cytokines to stimulate hematopoietic colony formation Martin et al. (1990) *Cell* 63, 203), stimulate hematopoiesis (Andrews et al. (1991) *Blood* 78, 1975), and rescue from the effects of lethal irradiation (Zsebo et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 9464) can be combined with the subject *RAFTK* therapeutics. When stem  
20 cells are induced to mature and/or proliferate *in vitro*, the subject *RAFTK* therapeutics can be combined with culture of the stem cells on feeder cells. Stem cells in which *RAFTK* is modulated can be useful, for example in the expansion of cells for autologous or allogeneic transplantation of stem cells or differentiated cells. *RAFTK* can be modulated to enhance engraftment and/or hematopoiesis after allogeneic bone  
25 marrow transplantation.

In preferred embodiments *RAFTK* can be modulated to control megakaryocyte development and to ameliorate diseases caused by abnormalities in megakaryocytic cells, for example, thrombocytopenia, myelodysplastic syndrome, myeloproliferative disorder, aplastic anemia, chronic myelogenous leukemia. Platelets are derived from  
30 megakaryocytes, and the subject *RAFTK* molecules can be used to correct abnormalities in platelet number or function. Platelets are important in numerous hemorrhagic and thrombotic disorders. In a preferred embodiment *RAFTK* bioactivity can be modulated to control platelet aggregation. In yet another embodiment, the subject *RAFTK* therapeutics can play a role in the maturation of cells of the erythroid  
35 lineage.

Among the approaches which can be used to ameliorate disease symptoms involving the misexpression of a *RAFTK* gene are, for example, antisense, ribozyme,

and triple helix molecules described above. Compounds that compete with an *RAFTK* protein for binding with an active portion of *RAFTK* will antagonize a *RAFTK* protein, thereby inducing a therapeutic effect. Examples of suitable compounds include the antagonists or homologues described in detail above. In other instances, the increased expression or bioactivity of a *RAFTK* protein may be desirable and may be accomplished by, for example the use of the *RAFTK* agonists or mimetics or by gene replacement therapy, as described herein.

#### *Effective Dose*

It is within the level of ordinary skill in the art to determine dosages of the subject *RAFTK* therapeutics. The data obtained from cell culture assays and animal studies can be used in formulating a range of dosage for use in humans. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED<sub>50</sub> with little or no toxicity. The dosage can vary within this range depending upon the dosage form employed and the route of administration utilized. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. A dose can be formulated in animal models to achieve a circulating plasma concentration range that includes the IC<sub>50</sub> (i.e., the concentration of the test compound which achieves a half-maximal inhibition of symptoms) as determined in cell culture. Such information can be used to more accurately determine useful doses in humans.

#### *Formulation and Use*

Pharmaceutical compositions for use in accordance with the present invention can be formulated in conventional manner using one or more physiologically acceptable carriers or excipients. Thus, the compounds (e.g., *RAFTK* polypeptides or *RAFTK* nucleic acids) and their physiologically acceptable salts and solvates may be formulated for administration by, for example, injection, inhalation or insufflation (either through the mouth or the nose) or oral, buccal, parenteral or rectal administration.

For such therapy, the oligomers of the invention can be formulated for a variety of loads of administration, including systemic and topical or localized administration. Techniques and formulations generally can be found in *Remington's Pharmaceutical Sciences*, Meade Publishing Co., Easton, PA. For systemic administration, injection is preferred, including intramuscular, intravenous, intraperitoneal, and subcutaneous.

For oral administration, the pharmaceutical compositions can take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g.,  
5 lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica); disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art.

Preparations for oral administration may be suitably formulated to give  
10 controlled release of the active compound.

The compounds can be formulated for parenteral administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions can take such forms as suspensions, solutions  
15 or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

Systemic administration can also be by transmucosal or transdermal means.  
20 For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration bile salts and fusidic acid derivatives. In addition, detergents may be used to facilitate permeation. Transmucosal administration may be through nasal sprays or using suppositories. For  
25 topical administration, the oligomers of the invention are formulated into ointments, salves, gels, or creams as generally known in the art.

In clinical settings, the gene delivery systems for the therapeutic *RAFTK* gene can be introduced into a patient by any of a number of methods, each of which is familiar in the art. For example, a pharmaceutical preparation of the gene delivery  
30 system can be introduced systemically, e.g. by intravenous injection, and specific transduction of the protein in the target cells occurs predominantly from specificity of transfection provided by the gene delivery vehicle, cell-type or tissue-type expression due to the transcriptional regulatory sequences controlling expression of the receptor gene, or a combination thereof. In other embodiments, initial delivery of the  
35 recombinant gene is localized. For example, the gene delivery vehicle can be introduced by catheter (see U.S. Patent 5,328,470) or by stereotactic injection (e.g. Chen et al. (1994) *PNAS* 91, 3054-3057). A mammalian *RAFTK* gene, such as any

one of the sequences represented in one of SEQ ID NOS:1 or 3, or a sequence homologous thereto can be delivered in a gene therapy construct by electroporation using techniques described, for example, by Dev et al. ((1994) *Cancer Treat Rev* 20, 105-115).

5           The pharmaceutical preparation of the gene therapy construct can include the gene delivery system in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery system can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can comprise one or more cells which produce  
10   the gene delivery system.

          The compositions can, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device can be accompanied by instructions for administration.

15

#### *Diagnostic and Prognostic Assays*

          The present method provides a method for determining if a subject is at risk for a disorder characterized by aberrant cell proliferation and/or differentiation. In preferred embodiments, the methods can be characterized as comprising detecting, in  
20   a sample of cells from the subject, the presence or absence of a genetic lesion characterized by at least one of (i) an alteration affecting the integrity of a gene encoding a *RAFTK*-protein, or (ii) the mis-expression of the *RAFTK* gene. To illustrate, such genetic lesions can be detected by ascertaining the existence of at least one of (i) a deletion of one or more nucleotides from a *RAFTK* gene, (ii) an addition  
25   of one or more nucleotides to a *RAFTK* gene, (iii) a substitution of one or more nucleotides of a *RAFTK* gene, (iv) a gross chromosomal rearrangement of a *RAFTK* gene, (v) a gross alteration in the level of a messenger RNA transcript of a *RAFTK* gene, (vii) aberrant modification of a *RAFTK* gene, such as of the methylation pattern of the genomic DNA, (vii) the presence of a non-wild type splicing pattern of a  
30   messenger RNA transcript of a *RAFTK* gene, (viii) a non-wild type level of a *RAFTK*-protein, (ix) allelic loss of a *RAFTK* gene, and (x) inappropriate post-translational modification of a *RAFTK*-protein. As described herein, the present invention provides a large number of assay techniques for detecting lesions in a *RAFTK* gene, and importantly, provides the ability to discern between different molecular causes  
35   underlying *RAFTK*-dependent aberrant cell growth, proliferation and/or differentiation.

In an exemplary embodiment, there is provided a nucleic acid composition comprising a (purified) oligonucleotide probe including a region of nucleotide sequence which is capable of hybridizing to a sense or antisense sequence of a *RAFTK* gene, such as represented by one of SEQ ID NOs:1 or 3, or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with the subject *RAFTK* genes or naturally occurring mutants thereof. The nucleic acid of a cell is rendered accessible for hybridization, the probe is exposed to nucleic acid of the sample, and the hybridization of the probe to the sample nucleic acid is detected. Such techniques can be used to detect lesions at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels.

In certain embodiments, detection of the lesion comprises utilizing the probe/primer in a polymerase chain reaction (PCR) (see, e.g. U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (see, e.g., Landegran et al. (1988) *Science* 241, 1077-1080; and Nakazawa et al. (1994) *PNAS* 91, 360-364), the latter of which can be particularly useful for detecting point mutations in the *RAFTK*-gene (see Abravaya et al. (1995) *Nuc Acid Res* 23, 675-682). In an illustrative embodiment, the method includes the steps of (i) collecting a sample of cells from a patient, (ii) isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, (iii) contacting the nucleic acid sample with one or more primers which specifically hybridize to a *RAFTK* gene under conditions such that hybridization and amplification of the *RAFTK*-gene (if present) occurs, and (iv) detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein. Alternative amplification methods include: self sustained sequence replication (Guatelli, J.C. et al. (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874-1878), transcriptional amplification system (Kwoh, D.Y. et al. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1173-1177), Q-Beta Replicase (Lizardi, P.M. et al. (1988) *Bio/Technology* 6, 1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In a preferred embodiment of the subject assay, mutations in a *RAFTK* gene from a sample cell are identified by alterations in restriction enzyme cleavage

patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis. Moreover, the use of sequence specific ribozymes (see, for example, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the *RAFTK* gene and detect mutations by comparing the sequence of the sample *RAFTK* with the corresponding wild-type (control) sequence. Exemplary sequencing reactions include those based on techniques developed by Maxim and Gilbert (*Proc. Natl Acad Sci USA* (1977) 74, 560) or Sanger (Sanger et al (1977) *Proc. Nat. Acad. Sci* 74, 5463). It is also contemplated that any of a variety of automated sequencing procedures may be utilized when performing the subject assays ((1995) *Biotechniques* 19, 448), including by sequencing by mass spectrometry (see, for example PCT publication WO 94/16101; Cohen et al. (1996) *Adv Chromatogr* 36, 127-162; and Griffin et al. (1993) *Appl Biochem Biotechnol* 38, 147-159). It will be evident to one skilled in the art that, for certain embodiments, the occurrence of only one, two or three of the nucleic acid bases need be determined in the sequencing reaction. For instance, A-tract or the like, e.g., where only one nucleic acid is detected, can be carried out.

In a further embodiment, protection from cleavage agents (such as a nuclease, hydroxylamine or osmium tetroxide and with piperidine) can be used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers, et al. (1985) *Science* 230, 1242). In general, the art technique of "mismatch cleavage" starts by providing heteroduplexes of formed by hybridizing (labelled) RNA or DNA containing the wild-type *RAFTK* sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to basepair mismatches between the control and sample strands. For instance, RNA/DNA duplexes can be treated with RNase and DNA/DNA hybrids treated with S1 nuclease to enzymatically digesting the mismatched regions. In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. See, for example, Cotton et al (1988) *Proc. Natl Acad Sci USA* 85, 4397; Saleeba et al (1992) *Methods Enzymol.* 217, 286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called "DNA mismatch repair" enzymes) in defined systems for detecting and mapping point mutations in *RAFTK* cDNAs obtained from samples of cells. For example, the mutY enzyme of *E. coli* cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al. (1994) *Carcinogenesis* 15, 1657-1662). According to an exemplary embodiment, a probe based on a *RAFTK* sequence, e.g., a wild-type *RAFTK* sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. See, for example, U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility is used to identify mutations in *RAFTK* genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) *Proc Natl. Acad. Sci USA* 86, 2766, see also Cotton (1993) *Mutat Res* 285, 125-144; and Hayashi (1992) *Genet Anal Tech Appl* 9, 73-79). Single-stranded DNA fragments of sample and control *RAFTK* nucleic acids is denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labelled or detected with labelled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) *Trends Genet* 7, 5).

In yet another embodiment the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al (1985) *Nature* 313, 495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing agent gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys Chem* 265, 12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or

selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al. (1986) *Nature* 324, 163); Saiki et al (1989) *Proc. Natl Acad. Sci USA* 86, 6230).

- 5 Such allele specific oligonucleotide hybridization techniques may be used to test one mutation per reaction when oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labelled target DNA.

- Alternatively, allele specific amplification technology which depends on  
10 selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al (1989) *Nucleic Acids Res.* 17, 2437-2448) or at the extreme  
15 3' end of one primer where, under appropriate conditions, mismatch can prevent, or reduce polymerase extension (Prossner (1993) *Tibtech* 11, 238. In addition it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al (1992) *Mol. Cell Probes* 6, 1). It is anticipated that in certain embodiments amplification may also be performed using  
20 Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci USA* 88, 189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

- Another embodiment of the invention provides for a nucleic acid composition comprising a (purified) oligonucleotide probe including a region of nucleotide  
25 sequence which is capable of hybridizing to a sense or antisense sequence of a *RAFTK*-gene, or naturally occurring mutants thereof, or 5' or 3' flanking sequences or intronic sequences naturally associated with the subject *RAFTK*-genes or naturally occurring mutants thereof. The nucleic acid of a cell is rendered accessible for hybridization, the probe is exposed to nucleic acid of the sample, and the  
30 hybridization of the probe to the sample nucleic acid is detected. Such techniques can be used to detect lesions at either the genomic or mRNA level, including deletions, substitutions, etc., as well as to determine mRNA transcript levels. Such oligonucleotide probes can be used for both predictive and therapeutic evaluation of allelic mutations which might be manifest in, for example, neoplastic or hyperplastic  
35 disorders (e.g. aberrant cell growth).

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody



reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose patients exhibiting symptoms or family history of a disease or illness involving a *RAFTK* gene.

Diagnostic procedures may also be performed *in situ* directly upon tissue  
5 sections (fixed and/or frozen) of patient tissue obtained from biopsies or resections, such that no nucleic acid purification is necessary. Nucleic acid reagents may be used as probes and/or primers for such *in situ* procedures (see, for example, Nuovo, G.J., 1992, *PCR in situ hybridization: protocols and applications*, Raven Press, NY).

In yet another embodiment mutant *RAFTK* proteins can be detected using the  
10 protein truncation test (PTT) (Dowton and Salugh. 1995, *Clin. Chem* 41:785). For PTT, RNA is initially isolated and reverse-transcribed, and the segment of interest is amplified by PCR. The PCR products are then used as a template for nested PCR amplification with a primer containing an RNA polymerase promoter and a translation initiation sequence. After amplification, the unique motifs incorporated into the  
15 primer permit sequential *in vitro* transcription and translation of the PCR products. Protein products are analyzed by electrophoresis and mutantations which cause truncation of the protein are identified by a change in the molecular weight of the protein. DNA may also be used.

In addition to methods which focus primarily on the detection of one nucleic  
20 acid sequence, profiles may also be assessed in such detection schemes. Fingerprint profiles may be generated, for example, by utilizing a differential display procedure, Northern analysis and/or RT-PCR.

Antibodies directed against wild type or mutant *RAFTK* proteins, which are discussed, above, may also be used indisease diagnostics and prognostics. Such  
25 diagnostic methods, may be used to detect abnormalities in the level of *RAFTK* protein expression, or abnormalities in the structure and/or tissue, cellular, or subcellular location of *RAFTK* protein. Structural differences may include, for example, differences in the size, electronegativity, or antigenicity of the mutant *RAFTK* protein relative to the normal *RAFTK* protein. Protein from the tissue or cell  
30 type to be analyzed may easily be detected or isolated using techniques which are well known to one of skill in the art, including but not limited to western blot analysis. For a detailed explanation of methods for carrying out western blot analysis, see Sambrook et al, 1989, *supra*, at Chapter 18. The protein detection and isolation methods employed herein may also be such as those described in Harlow and Lane,  
35 for example, (Harlow, E. and Lane, D., 1988, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York), which is incorporated herein by reference in its entirety.

This can be accomplished, for example, by immunofluorescence techniques employing a fluorescently labeled antibody (see below) coupled with light microscopic, flow cytometric, or fluorimetric detection. The antibodies (or fragments thereof) useful in the present invention may, additionally, be employed histologically, as in immunofluorescence or immunoelectron microscopy, for in situ detection of *RAFTK* proteins. In situ detection may be accomplished by removing a histological specimen from a patient, and contacting it with a labeled antibody of the present invention. The antibody (or fragment) is preferably applied by overlaying the labeled antibody (or fragment) onto a biological sample. Through the use of such a procedure, it is possible to determine not only the presence of the *RAFTK* protein, but also its distribution in the examined tissue. Using the present invention, one of ordinary skill will readily perceive that any of a wide variety of histological methods (such as staining procedures) can be modified in order to achieve such in situ detection.

Often a solid phase support or carrier is used as a support capable of binding an antigen or an antibody. Well-known supports or carriers include glass, polystyrene, polypropylene, polyethylene, dextran, nylon, amylases, natural and modified celluloses, polyacrylamides, gabbros, and magnetite. The nature of the carrier can be either soluble to some extent or insoluble for the purposes of the present invention. The support material may have virtually any possible structural configuration so long as the coupled molecule is capable of binding to an antigen or antibody. Thus, the support configuration may be spherical, as in a bead, or cylindrical, as in the inside surface of a test tube, or the external surface of a rod. Alternatively, the surface may be flat such as a sheet, test strip, etc. Preferred supports include polystyrene beads. Those skilled in the art will know many other suitable carriers for binding antibody or antigen, or will be able to ascertain the same by use of routine experimentation.

One means for labeling an anti- *RAFTK* protein specific antibody is via linkage to an enzyme and use in an enzyme immunoassay (EIA) (Voller (1978), "The Enzyme Linked Immunosorbent Assay (ELISA)", *Diagnostic Horizons* 2, 1-7, 1978, Microbiological Associates Quarterly Publication, Walkersville, MD; Voller, et al., J. Clin. Pathol. 31, 507-520 (1978); Butler (1981) *Meth. Enzymol.* 73, 482-523; Maggio, (ed.) *Enzyme Immunoassay*, CRC Press, Boca Raton, FL, 1980; Ishikawa, et al., (eds.) *Enzyme Immunoassay*, Kigaku Shoin, Tokyo, 1981). The enzyme which is bound to the antibody will react with an appropriate substrate, preferably a chromogenic substrate, in such a manner as to produce a chemical moiety which can be detected, for example, by spectrophotometric, fluorimetric or by visual means.

Enzymes which can be used to detectably label the antibody include, but are not limited to, malate dehydrogenase, staphylococcal nuclease, delta-5-steroid isomerase, yeast alcohol dehydrogenase, alpha-glycerophosphate, dehydrogenase, triose phosphate isomerase, horseradish peroxidase, alkaline phosphatase, asparaginase, glucose oxidase, beta-galactosidase, ribonuclease, urease, catalase, glucose-6-phosphate dehydrogenase, glucoamylase and acetylcholinesterase. The detection can be accomplished by colorimetric methods which employ a chromogenic substrate for the enzyme. Detection may also be accomplished by visual comparison of the extent of enzymatic reaction of a substrate in comparison with similarly prepared standards.

Detection may also be accomplished using any of a variety of other immunoassays. For example, by radioactively labeling the antibodies or antibody fragments, it is possible to detect fingerprint gene wild type or mutant peptides through the use of a radioimmunoassay (RIA) (see, for example, Weintraub, B., *Principles of Radioimmunoassays*, Seventh Training Course on Radioligand Assay Techniques, The Endocrine Society, March, 1986, which is incorporated by reference herein). The radioactive isotope can be detected by such means as the use of a gamma counter or a scintillation counter or by autoradiography.

It is also possible to label the antibody with a fluorescent compound. When the fluorescently labeled antibody is exposed to light of the proper wave length, its presence can then be detected due to fluorescence. Among the most commonly used fluorescent labeling compounds are fluorescein isothiocyanate, rhodamine, phycoerythrin, phycocyanin, allophycocyanin, o-phthaldehyde and fluorescamine. The antibody can also be labeled using fluorescence emitting metals such as <sup>152</sup>Eu, or others of the lanthanide series. These metals can be attached to the antibody using such metal chelating groups as diethylenetriaminepentacetic acid (DTPA) or ethylenediaminetetraacetic acid (EDTA).

The antibody also can be detectably labeled by coupling it to a chemiluminescent compound. The presence of the chemiluminescent-tagged antibody is then determined by detecting the presence of luminescence that arises during the course of a chemical reaction. Examples of particularly useful chemiluminescent labeling compounds are luminol, isoluminol, thermotropic acridinium ester, imidazole, acridinium salt and oxalate ester. Likewise, a bioluminescent compound may be used to label the antibody of the present invention. Bioluminescence is a type of chemiluminescence found in biological systems in which a catalytic protein increases the efficiency of the chemiluminescent reaction. The presence of a bioluminescent protein is determined by detecting luminescence. Important bioluminescent compounds for purposes of labeling are luciferin, luciferase and aequorin.

Moreover, it will be understood that any of the above methods for detecting alterations in a *RAFTK* gene or gene product can be used to monitor the course of treatment or therapy.

## 5 *Drug Screening Assays*

Furthermore, by making available purified and recombinant *RAFTK* polypeptides, the present invention facilitates the development of assays which can be used to screen for compounds, including *RAFTK* homologs, which are either agonists or antagonists of the normal cellular function of the subject *RAFTK* polypeptides, or  
10 of their role in the pathogenesis of cellular differentiation and/or proliferation and disorders related thereto. A variety of assay formats can be utilized and, in light of the present inventions, will be comprehended by a skilled artisan.

### *Cell-free assays*

15 In many drug screening programs which test libraries of compounds and natural extracts, high throughput assays are desirable in order to maximize the number of compounds surveyed in a given period of time. Assays which are performed in cell-free systems, such as may be derived with purified or semi-purified proteins, are often preferred as "primary" screens in that they can be generated to permit rapid  
20 development and relatively easy detection of an alteration in a molecular target which is mediated by a test compound. Moreover, the effects of cellular toxicity and/or bioavailability of the test compound can be generally ignored in the in vitro system, the assay instead can be focused primarily on the effect of the drug on the molecular target as may be manifest in an alteration of binding affinity with upstream or  
25 downstream elements. Accordingly, in an exemplary screening assay of the present invention, the compound of interest is contacted with proteins which may function upstream (including both activators and repressors of its activity) or to proteins or nucleic acids which may function downstream of the *RAFTK* polypeptide, whether they are positively or negatively regulated by it. To the mixture of the compound and  
30 the upstream or downstream element is then added a composition containing a *RAFTK* polypeptide. Detection and quantification of complexes of *RAFTK* with its upstream or downstream elements provide a means for determining a compound's efficacy at inhibiting (or potentiating) complex formation between *RAFTK* and the *RAFTK*-binding elements. The efficacy of the compound can be assessed by  
35 generating dose response curves from data obtained using various concentrations of the test compound. Moreover, a control assay can also be performed to provide a baseline for comparison. In the control assay, isolated and purified *RAFTK*

polypeptide is added to a composition containing the *RAFTK*-binding element, and the formation of a complex is quantitated in the absence of the test compound.

Complex formation between the *RAFTK* polypeptide and a *RAFTK* binding element may be detected by a variety of techniques. Modulation of the formation of  
5 complexes can be quantitated using, for example, detectably labeled proteins such as radiolabeled, fluorescently labeled, or enzymatically labeled *RAFTK* polypeptides, by immunoassay, or by chromatographic detection.

Typically, it is desirable to immobilize either *RAFTK* or its binding protein to facilitate separation of complexes from uncomplexed forms of one or both of the  
10 proteins, as well as to accommodate automation of the assay. Binding of *RAFTK* to an upstream or downstream element, in the presence and absence of a candidate agent, can be accomplished in any vessel suitable for containing the reactants. Examples include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows the protein to be  
15 bound to a matrix. For example, glutathione-S-transferase/*RAFTK* (GST/*RAFTK*) fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical, St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the cell lysates (e.g. <sup>35</sup>S-labeled) and the test compound, and the mixture incubated under conditions conducive to complex formation, for example at  
20 physiological conditions for salt and pH, though slightly more stringent conditions may be desired. Following incubation, the beads are washed to remove any unbound label, and the matrix immobilized and radiolabel determined directly (e.g. beads placed in scintillant), or in the supernatant after the complexes are dissociated. Alternatively, the complexes can be dissociated from the matrix, separated by SDS-  
25 PAGE, and the level of *RAFTK*-binding protein found in the bead fraction quantitated from the gel using standard electrophoretic techniques.

Other techniques for immobilizing proteins on matrices are also available for use in the subject assay. For instance, either *RAFTK* or its cognate binding protein can be immobilized utilizing conjugation of biotin and streptavidin. For instance,  
30 biotinylated *RAFTK* molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals, Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with *RAFTK* but which do not interfere with binding of upstream or downstream elements can be derivatized  
35 to the wells of the plate, and *RAFTK* trapped in the wells by antibody conjugation. As above, preparations of a *RAFTK*-binding protein and a test compound are incubated in the *RAFTK*-presenting wells of the plate, and the amount of complex trapped in the

well can be quantitated. Exemplary methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the *RAFTK* binding element, or which are reactive with *RAFTK* protein and compete with the binding  
5 element; as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the binding element, either intrinsic or extrinsic activity. In the instance of the latter, the enzyme can be chemically conjugated or provided as a fusion protein with the *RAFTK*-BP. To illustrate, the *RAFTK*-BP can be chemically cross-linked or genetically fused with horseradish peroxidase, and the amount of  
10 polypeptide trapped in the complex can be assessed with a chromogenic substrate of the enzyme, e.g. 3,3'-diamino-benzadine tetrahydrochloride or 4-chloro-1-naphthol. Likewise, a fusion protein comprising the polypeptide and glutathione-S-transferase can be provided, and complex formation quantitated by detecting the GST activity using 1-chloro-2,4-dinitrobenzene (Habig et al (1974) *J Biol Chem* 249, 7130).

15 For processes which rely on immunodetection for quantitating one of the proteins trapped in the complex, antibodies against the protein, such as anti-*RAFTK* antibodies, can be used. Alternatively, the protein to be detected in the complex can be "epitope tagged" in the form of a fusion protein which includes, in addition to the *RAFTK* sequence, a second polypeptide for which antibodies are readily available  
20 (e.g. from commercial sources). For instance, the GST fusion proteins described above can also be used for quantification of binding using antibodies against the GST moiety. Other useful epitope tags include myc-epitopes (e.g., see Ellison et al. (1991) *J Biol Chem* 266, 21150-21157) which includes a 10-residue sequence from c-myc, as well as the pFLAG system (International Biotechnologies, Inc.) or the pEZZ-protein  
25 A system (Pharmacia, NJ).

#### *Cell based assays*

In addition to cell-free assays, such as described above, the readily available source of mammalian *RAFTK* proteins provided by the present invention also  
30 facilitates the generation of cell-based assays for identifying small molecule agonists/antagonists and the like. For example, cells which are sensitive to ligands which transduce signals via a pathway involving *RAFTK* can be caused to overexpress a recombinant *RAFTK* protein in the presence and absence of a test agent of interest, with the assay scoring for modulation of *RAFTK* responses by the target cell  
35 mediated by the test agent. As with the cell-free assays, agents which produce a statistically significant change in *RAFTK*-dependent responses (either inhibition or potentiation) can be identified. In an illustrative embodiment, the expression or

activity of a *RAFTK* is modulated in cells and the effects of compounds of interest on the readout of interest (such as apoptosis, proliferation or differentiation) are measured. For example, the expression of genes which are up- or down-regulated in response to a *RAFTK*-dependent signal cascade can be assayed. In preferred  
5 embodiments, the regulatory regions of such genes, e.g., the 5' flanking promoter and enhancer regions, are operably linked to a detectable marker (such as luciferase) which encodes a gene product that can be readily detected. Phosphorylation of *RAFTK* or *RAFTK* binding proteins can also be measured, for example by immunoblotting as described in the appended examples.

10 Monitoring the influence of compounds on cells may be applied not only in basic drug screening, but also in clinical trials. In such clinical trials, the expression of a panel of genes may be used as a "read out" of a particular drug's therapeutic effect.

In yet another aspect of the invention, the subject *RAFTK* polypeptides can be  
15 used to generate a "two hybrid" assay (see, for example, U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72, 223-232; Madura et al. (1993) *J Biol Chem* 268, 12046-12054; Bartel et al. (1993) *Biotechniques* 14, 920-924; Iwabuchi et al. (1993) *Oncogene* 8, 1693-1696; and Brent WO94/10300), for isolating coding sequences for other cellular proteins which bind to or interact with *RAFTK* ("*RAFTK*-binding  
20 proteins" or "*RAFTK*-bp"). Such *RAFTK*-binding proteins would likely also be involved in the propagation of signals by the *RAFTK* proteins as, for example, the upstream or downstream elements of the *RAFTK* pathway.

Briefly, the two hybrid assay relies on reconstituting in vivo a functional transcriptional activator protein from two separate fusion proteins. In particular, the  
25 method makes use of chimeric genes which express hybrid proteins. To illustrate, a first hybrid gene comprises the coding sequence for a DNA-binding domain of a transcriptional activator fused in frame to the coding sequence for a *RAFTK* polypeptide. The second hybrid protein encodes a transcriptional activation domain fused in frame to a sample gene from a cDNA library. If the bait and sample hybrid  
30 proteins are able to interact, e.g., form a *RAFTK*-dependent complex, they bring into close proximity the two domains of the transcriptional activator. This proximity is sufficient to cause transcription of a reporter gene which is operably linked to a transcriptional regulatory site responsive to the transcriptional activator, and expression of the reporter gene can be detected and used to score for the interaction of  
35 the *RAFTK* and sample proteins.

*Transgenic animals*

These systems may be used in a variety of applications. For example, the cell- and animal-based model systems may be used to further characterize *RAFTK* genes and proteins. In addition, such assays may be utilized as part of screening strategies designed to identify compounds which are capable of ameliorating disease symptoms. Thus, the animal- and cell-based models may be used to identify drugs, pharmaceuticals, therapies and interventions which may be effective in treating disease.

One aspect of the present invention concerns transgenic animals comprising of cells which contain a transgene of the present invention and which preferably (though optionally) express an exogenous *RAFTK* protein in one or more cells in the animal. A *RAFTK* transgene can encode the wild-type form of the protein, or can encode homologs thereof, including both agonists and antagonists, as well as antisense constructs. In preferred embodiments, the expression of the transgene is restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, cis-acting sequences that control expression in the desired pattern. In the present invention, such mosaic expression of a *RAFTK* protein can be essential for many forms of lineage analysis and can additionally provide a means to assess the effects of, for example, lack of *RAFTK* expression which might grossly alter development in small patches of tissue within an otherwise normal embryo. Toward this end, tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory sequences.

Genetic techniques which allow for the expression of transgenes can be regulated via site-specific genetic manipulation in vivo are known to those skilled in the art. For instance, genetic systems are available which allow for the regulated expression of a recombinase that catalyzes the genetic recombination at a target sequence. As used herein, the phrase "target sequence" refers to a nucleotide sequence that is genetically recombined by a recombinase. The target sequence is flanked by recombinase recognition sequences and is generally either excised or inverted in cells expressing recombinase activity. Recombinase catalyzed recombination events can be designed such that recombination of the target sequence results in either the activation or repression of expression of one of the subject *RAFTK* proteins. For example, excision of a target sequence which interferes with the expression of a recombinant *RAFTK* gene, such as one which encodes an antagonistic homolog or an antisense transcript, can be designed to activate expression of that



gene. This interference with expression of the protein can result from a variety of mechanisms, such as spatial separation of the *RAFTK* gene from the promoter element or an internal stop codon. Moreover, the transgene can be made wherein the coding sequence of the gene is flanked by recombinase recognition sequences and is initially  
5 transfected into cells in a 3' to 5' orientation with respect to the promoter element. In such an instance, inversion of the target sequence will reorient the subject gene by placing the 5' end of the coding sequence in an orientation with respect to the promoter element which allow for promoter driven transcriptional activation.

The transgenic animals of the present invention all include within a plurality  
10 of their cells a transgene of the present invention, which transgene alters the phenotype of the "host cell" with respect to regulation of cell growth, death and/or differentiation. Since it is possible to produce transgenic organisms of the invention utilizing one or more of the transgene constructs described herein, a general description will be given of the production of transgenic organisms by referring  
15 generally to exogenous genetic material. This general description can be adapted by those skilled in the art in order to incorporate specific transgene sequences into organisms utilizing the methods and materials described below.

In an illustrative embodiment, either the cre/loxP recombinase system of bacteriophage P1 (Lakso et al. (1992) *PNAS* 89, 6232-6236; Orban et al. (1992) *PNAS*  
20 89, 6861-6865) or the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) *Science* 251, 1351-1355; PCT publication WO 92/15694) can be used to generate in vivo site-specific genetic recombination systems. Cre recombinase catalyzes the site-specific recombination of an intervening target sequence located between loxP sequences. loxP sequences are 34 base pair nucleotide  
25 repeat sequences to which the Cre recombinase binds and are required for Cre recombinase mediated genetic recombination. The orientation of loxP sequences determines whether the intervening target sequence is excised or inverted when Cre recombinase is present (Abremski et al. (1984) *J. Biol. Chem.* 259, 1509-1514); catalyzing the excision of the target sequence when the loxP sequences are oriented as  
30 direct repeats and catalyzes inversion of the target sequence when loxP sequences are oriented as inverted repeats.

Accordingly, genetic recombination of the target sequence is dependent on expression of the Cre recombinase. Expression of the recombinase can be regulated by promoter elements which are subject to regulatory control, e.g., tissue-specific,  
35 developmental stage-specific, inducible or repressible by externally added agents. This regulated control will result in genetic recombination of the target sequence only in cells where recombinase expression is mediated by the promoter element. Thus,

the activation expression of a recombinant *RAFTK* protein can be regulated via control of recombinase expression.

Use of the cre/loxP recombinase system to regulate expression of a recombinant *RAFTK* protein requires the construction of a transgenic animal  
5 containing transgenes encoding both the Cre recombinase and the subject protein. Animals containing both the Cre recombinase and a recombinant *RAFTK* gene can be provided through the construction of "double" transgenic animals. A convenient method for providing such animals is to mate two transgenic animals each containing a transgene, e.g., a *RAFTK* gene and recombinase gene.

10 One advantage derived from initially constructing transgenic animals containing a *RAFTK* transgene in a recombinase-mediated expressible format derives from the likelihood that the subject protein, whether agonistic or antagonistic, can be deleterious upon expression in the transgenic animal. In such an instance, a founder population, in which the subject transgene is silent in all tissues, can be propagated  
15 and maintained. Individuals of this founder population can be crossed with animals expressing the recombinase in, for example, one or more tissues and/or a desired temporal pattern. Thus, the creation of a founder population in which, for example, an antagonistic *RAFTK* transgene is silent will allow the study of progeny from that founder in which disruption of *RAFTK* mediated induction in a particular tissue or at  
20 certain developmental stages would result in, for example, a lethal phenotype.

Similar conditional transgenes can be provided using prokaryotic promoter sequences which require prokaryotic proteins to be simultaneous expressed in order to facilitate expression of the *RAFTK* transgene. Exemplary promoters and the corresponding trans-activating prokaryotic proteins are given in U.S. Patent No.  
25 4,833,080.

Moreover, expression of the conditional transgenes can be induced by gene therapy-like methods wherein a gene encoding the trans-activating protein, e.g. a recombinase or a prokaryotic protein, is delivered to the tissue and caused to be expressed, such as in a cell-type specific manner. By this method, a *RAFTK* transgene  
30 could remain silent into adulthood until "turned on" by the introduction of the trans-activator.

In an exemplary embodiment, the "transgenic non-human animals" of the invention are produced by introducing transgenes into the germline of the non-human animal. Embryonal target cells at various developmental stages can be used to  
35 introduce transgenes. Different methods are used depending on the stage of development of the embryonal target cell. The specific line(s) of any animal used to practice this invention are selected for general good health, good embryo yields, good

pronuclear visibility in the embryo, and good reproductive fitness. In addition, the haplotype is a significant factor. For example, when transgenic mice are to be produced, strains such as C57BL/6 or FVB lines are often used (Jackson Laboratory, Bar Harbor, ME). Preferred strains are those with H-2b, H-2d or H-2q haplotypes such as C57BL/6 or DBA/1. The line(s) used to practice this invention may themselves be transgenics, and/or may be knockouts (i.e., obtained from animals which have one or more genes partially or completely suppressed) .

In one embodiment, the transgene construct is introduced into a single stage embryo. The zygote is the best target for micro-injection. In the mouse, the male pronucleus reaches the size of approximately 20 micrometers in diameter which allows reproducible injection of 1-2pl of DNA solution. The use of zygotes as a target for gene transfer has a major advantage in that in most cases the injected DNA will be incorporated into the host gene before the first cleavage (Brinster et al. (1985) *PNAS* 82, 4438-4442). As a consequence, all cells of the transgenic animal will carry the incorporated transgene. This will in general also be reflected in the efficient transmission of the transgene to offspring of the founder since 50% of the germ cells will harbor the transgene.

Normally, fertilized embryos are incubated in suitable media until the pronuclei appear. At about this time, the nucleotide sequence comprising the transgene is introduced into the female or male pronucleus as described below. In some species such as mice, the male pronucleus is preferred. It is most preferred that the exogenous genetic material be added to the male DNA complement of the zygote prior to its being processed by the ovum nucleus or the zygote female pronucleus. It is thought that the ovum nucleus or female pronucleus release molecules which affect the male DNA complement, perhaps by replacing the protamines of the male DNA with histones, thereby facilitating the combination of the female and male DNA complements to form the diploid zygote.

Thus, it is preferred that the exogenous genetic material be added to the male complement of DNA or any other complement of DNA prior to its being affected by the female pronucleus. For example, the exogenous genetic material is added to the early male pronucleus, as soon as possible after the formation of the male pronucleus, which is when the male and female pronuclei are well separated and both are located close to the cell membrane. Alternatively, the exogenous genetic material could be added to the nucleus of the sperm after it has been induced to undergo decondensation. Sperm containing the exogenous genetic material can then be added to the ovum or the decondensed sperm could be added to the ovum with the transgene constructs being added as soon as possible thereafter.

Introduction of the transgene nucleotide sequence into the embryo may be accomplished by any means known in the art such as, for example, microinjection, electroporation, or lipofection. Following introduction of the transgene nucleotide sequence into the embryo, the embryo may be incubated in vitro for varying amounts  
5 of time, or reimplanted into the surrogate host, or both. In vitro incubation to maturity is within the scope of this invention. One common method in to incubate the embryos in vitro for about 1-7 days, depending on the species, and then reimplant them into the surrogate host.

For the purposes of this invention a zygote is essentially the formation of a  
10 diploid cell which is capable of developing into a complete organism. Generally, the zygote will be comprised of an egg containing a nucleus formed, either naturally or artificially, by the fusion of two haploid nuclei from a gamete or gametes. Thus, the gamete nuclei must be ones which are naturally compatible, i.e., ones which result in a viable zygote capable of undergoing differentiation and developing into a functioning  
15 organism. Generally, a euploid zygote is preferred. If an aneuploid zygote is obtained, then the number of chromosomes should not vary by more than one with respect to the euploid number of the organism from which either gamete originated.

In addition to similar biological considerations, physical ones also govern the amount (e.g., volume) of exogenous genetic material which can be added to the  
20 nucleus of the zygote or to the genetic material which forms a part of the zygote nucleus. If no genetic material is removed, then the amount of exogenous genetic material which can be added is limited by the amount which will be absorbed without being physically disruptive. Generally, the volume of exogenous genetic material inserted will not exceed about 10 picoliters. The physical effects of addition must not  
25 be so great as to physically destroy the viability of the zygote. The biological limit of the number and variety of DNA sequences will vary depending upon the particular zygote and functions of the exogenous genetic material and will be readily apparent to one skilled in the art, because the genetic material, including the exogenous genetic material, of the resulting zygote must be biologically capable of initiating and  
30 maintaining the differentiation and development of the zygote into a functional organism.

The number of copies of the transgene constructs which are added to the zygote is dependent upon the total amount of exogenous genetic material added and will be the amount which enables the genetic transformation to occur. Theoretically  
35 only one copy is required; however, generally, numerous copies are utilized, for example, 1,000-20,000 copies of the transgene construct, in order to insure that one copy is functional. As regards the present invention, there will often be an advantage

to having more than one functioning copy of each of the inserted exogenous DNA sequences to enhance the phenotypic expression of the exogenous DNA sequences.

Any technique which allows for the addition of the exogenous genetic material into nucleic genetic material can be utilized so long as it is not destructive to the cell, nuclear membrane or other existing cellular or genetic structures. The exogenous genetic material is preferentially inserted into the nucleic genetic material by microinjection. Microinjection of cells and cellular structures is known and is used in the art.

Reimplantation is accomplished using standard methods. Usually, the surrogate host is anesthetized, and the embryos are inserted into the oviduct. The number of embryos implanted into a particular host will vary by species, but will usually be comparable to the number of off spring the species naturally produces.

Transgenic offspring of the surrogate host may be screened for the presence and/or expression of the transgene by any suitable method. Screening is often accomplished by Southern blot or Northern blot analysis, using a probe that is complementary to at least a portion of the transgene. Western blot analysis using an antibody against the protein encoded by the transgene may be employed as an alternative or additional method for screening for the presence of the transgene product. Typically, DNA is prepared from tail tissue and analyzed by Southern analysis or PCR for the transgene. Alternatively, the tissues or cells believed to express the transgene at the highest levels are tested for the presence and expression of the transgene using Southern analysis or PCR, although any tissues or cell types may be used for this analysis.

Alternative or additional methods for evaluating the presence of the transgene include, without limitation, suitable biochemical assays such as enzyme and/or immunological assays, histological stains for particular marker or enzyme activities, flow cytometric analysis, and the like. Analysis of the blood may also be useful to detect the presence of the transgene product in the blood, as well as to evaluate the effect of the transgene on the levels of various types of blood cells and other blood constituents.

Progeny of the transgenic animals may be obtained by mating the transgenic animal with a suitable partner, or by *in vitro* fertilization of eggs and/or sperm obtained from the transgenic animal. Where mating with a partner is to be performed, the partner may or may not be transgenic and/or a knockout; where it is transgenic, it may contain the same or a different transgene, or both. Alternatively, the partner may be a parental line. Where *in vitro* fertilization is used, the fertilized embryo may be

implanted into a surrogate host or incubated in vitro, or both. Using either method, the progeny may be evaluated for the presence of the transgene using methods described above, or other appropriate methods.

5 The transgenic animals produced in accordance with the present invention will include exogenous genetic material. As set out above, the exogenous genetic material will, in certain embodiments, be a DNA sequence which results in the production of a *RAFTK* protein (either agonistic or antagonistic), and antisense transcript, or a *RAFTK* mutant. Further, in such embodiments the sequence will be attached to a transcriptional control element, e.g., a promoter, which preferably allows the  
10 expression of the transgene product in a specific type of cell.

Retroviral infection can also be used to introduce transgene into a non-human animal. The developing non-human embryo can be cultured in vitro to the blastocyst stage. During this time, the blastomeres can be targets for retroviral infection (Jaenich, R. (1976) *PNAS* 73, 1260-1264). Efficient infection of the blastomeres is obtained by  
15 enzymatic treatment to remove the zona pellucida (*Manipulating the Mouse Embryo*, Hogan eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1986). The viral vector system used to introduce the transgene is typically a replication-defective retrovirus carrying the transgene (Jahner et al. (1985) *PNAS* 82, 6927-6931; Van der Putten et al. (1985) *PNAS* 82, 6148-6152). Transfection is easily and efficiently  
20 obtained by culturing the blastomeres on a monolayer of virus-producing cells (Van der Putten, supra; Stewart et al. (1987) *EMBO J.* 6, 383-388). Alternatively, infection can be performed at a later stage. Virus or virus-producing cells can be injected into the blastocoele (Jahner et al. (1982) *Nature* 298, 623-628). Most of the founders will be mosaic for the transgene since incorporation occurs only in a subset of the cells  
25 which formed the transgenic non-human animal. Further, the founder may contain various retroviral insertions of the transgene at different positions in the genome which generally will segregate in the offspring. In addition, it is also possible to introduce transgenes into the germ line by intrauterine retroviral infection of the midgestation embryo (Jahner et al. (1982) supra).

30 A third type of target cell for transgene introduction is the embryonal stem cell (ES). ES cells are obtained from pre-implantation embryos cultured in vitro and fused with embryos (Evans et al. (1981) *Nature* 292, 154-156; Bradley et al. (1984) *Nature* 309, 255-258; Gossler et al. (1986) *PNAS* 83, 9065-9069; and Robertson et al. (1986) *Nature* 322, 445-448). Transgenes can be efficiently introduced into the ES cells by  
35 DNA transfection or by retrovirus-mediated transduction. Such transformed ES cells can thereafter be combined with blastocysts from a non-human animal. The ES cells

thereafter colonize the embryo and contribute to the germ line of the resulting chimeric animal. For review see Jaenisch, R. (1988) *Science* 240, 1468-1474.

In one embodiment, gene targeting, which is a method of using homologous recombination to modify an animal's genome, can be used to introduce changes into cultured embryonic stem cells. By targeting a *RAFTK* gene of interest in ES cells, these changes can be introduced into the germlines of animals to generate chimeras. The gene targeting procedure is accomplished by introducing into tissue culture cells a DNA targeting construct that includes a segment homologous to a target *RAFTK* locus, and which also includes an intended sequence modification to the *RAFTK* genomic sequence (e.g., insertion, deletion, point mutation). The treated cells are then screened for accurate targeting to identify and isolate those which have been properly targeted.

Gene targeting in embryonic stem cells is in fact a scheme contemplated by the present invention as a means for disrupting a *RAFTK* gene function through the use of a targeting transgene construct designed to undergo homologous recombination with one or more *RAFTK* genomic sequences. The targeting construct can be arranged so that, upon recombination with an element of a *RAFTK* gene, a positive selection marker is inserted into (or replaces) coding sequences of the targeted *signanlin* gene. The inserted sequence functionally disrupts the *RAFTK* gene, while also providing a positive selection trait. Exemplary *RAFTK* targeting constructs are described in more detail below.

Generally, the embryonic stem cells (ES cells) used to produce the knockout animals will be of the same species as the knockout animal to be generated. Thus for example, mouse embryonic stem cells will usually be used for generation of knockout mice.

Embryonic stem cells are generated and maintained using methods well known to the skilled artisan such as those described by Doetschman et al. (1985) *J. Embryol. Exp. Morphol.* 87, 27-45). Any line of ES cells can be used, however, the line chosen is typically selected for the ability of the cells to integrate into and become part of the germ line of a developing embryo so as to create germ line transmission of the knockout construct. Thus, any ES cell line that is believed to have this capability is suitable for use herein. One mouse strain that is typically used for production of ES cells, is the 129J strain. Another ES cell line is murine cell line D3 (American Type Culture Collection, catalog no. CKL 1934) Still another preferred ES cell line is the WW6 cell line (Ioffe et al. (1995) *PNAS* 92, 7357-7361). The cells are cultured and prepared for knockout construct insertion using methods well known to the skilled artisan, such as those set forth by Robertson in: *Teratocarcinomas and Embryonic*

*Stem Cells: A Practical Approach*, E.J. Robertson, ed. IRL Press, Washington, D.C. [1987]); by Bradley et al. (1986) *Current Topics in Devel. Biol.* 20, 357-371); and by Hogan et al. (*Manipulating the Mouse Embryo: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY [1986]) .

5           Insertion of the knockout construct into the ES cells can be accomplished using a variety of methods well known in the art including for example, electroporation, microinjection, and calcium phosphate treatment. A preferred method of insertion is electroporation .

10           Each knockout construct to be inserted into the cell must first be in the linear form. Therefore, if the knockout construct has been inserted into a vector (described infra), linearization is accomplished by digesting the DNA with a suitable restriction endonuclease selected to cut only within the vector sequence and not within the knockout construct sequence.

15           For insertion, the knockout construct is added to the ES cells under appropriate conditions for the insertion method chosen, as is known to the skilled artisan. Where more than one construct is to be introduced into the ES cell, each knockout construct can be introduced simultaneously or one at a time.

20           If the ES cells are to be electroporated, the ES cells and knockout construct DNA are exposed to an electric pulse using an electroporation machine and following the manufacturer's guidelines for use. After electroporation, the ES cells are typically allowed to recover under suitable incubation conditions. The cells are then screened for the presence of the knockout construct.

25           Screening can be accomplished using a variety of methods. Where the marker gene is an antibiotic resistance gene, for example, the ES cells may be cultured in the presence of an otherwise lethal concentration of antibiotic. Those ES cells that survive have presumably integrated the knockout construct. If the marker gene is other than an antibiotic resistance gene, a Southern blot of the ES cell genomic DNA can be probed with a sequence of DNA designed to hybridize only to the marker sequence. Alternatively, PCR can be used. Finally, if the marker gene is a gene that encodes an enzyme whose activity can be detected (e.g., b-galactosidase), the enzyme substrate  
30           can be added to the cells under suitable conditions, and the enzymatic activity can be analyzed. One skilled in the art will be familiar with other useful markers and the means for detecting their presence in a given cell. All such markers are contemplated as being included within the scope of the teaching of this invention.

35           The knockout construct may integrate into several locations in the ES cell genome, and may integrate into a different location in each ES cell's genome due to the occurrence of random insertion events. The desired location of insertion is in a



complementary position to the DNA sequence to be knocked out, e.g., the *RAFTK* coding sequence, transcriptional regulatory sequence, etc. Typically, less than about 1-5 % of the ES cells that take up the knockout construct will actually integrate the knockout construct in the desired location. To identify those ES cells with proper  
5 integration of the knockout construct, total DNA can be extracted from the ES cells using standard methods. The DNA can then be probed on a Southern blot with a probe or probes designed to hybridize in a specific pattern to genomic DNA digested with particular restriction enzyme(s). Alternatively, or additionally, the genomic DNA can be amplified by PCR with probes specifically designed to amplify DNA  
10 fragments of a particular size and sequence (i.e., only those cells containing the knockout construct in the proper position will generate DNA fragments of the proper size).

After suitable ES cells containing the knockout construct in the proper location have been identified, the cells can be inserted into an embryo. Insertion may be  
15 accomplished in a variety of ways known to the skilled artisan, however a preferred method is by microinjection. For microinjection, about 10-30 cells are collected into a micropipet and injected into embryos that are at the proper stage of development to permit integration of the foreign ES cell containing the knockout construct into the developing embryo. For instance, as the appended Examples describe, the  
20 transformed ES cells can be microinjected into blastocytes.

The suitable stage of development for the embryo used for insertion of ES cells is very species dependent, however for mice it is about 3.5 days. The embryos are obtained by perfusing the uterus of pregnant females. Suitable methods for accomplishing this are known to the skilled artisan, and are set forth by, e.g., Bradley  
25 et al. (*supra*).

While any embryo of the right stage of development is suitable for use, preferred embryos are male. In mice, the preferred embryos also have genes coding for a coat color that is different from the coat color encoded by the ES cell genes. In this way, the offspring can be screened easily for the presence of the knockout  
30 construct by looking for mosaic coat color (indicating that the ES cell was incorporated into the developing embryo). Thus, for example, if the ES cell line carries the genes for white fur, the embryo selected will carry genes for black or brown fur.

After the ES cell has been introduced into the embryo, the embryo may be  
35 implanted into the uterus of a pseudopregnant foster mother for gestation. While any foster mother may be used, the foster mother is typically selected for her ability to breed and reproduce well, and for her ability to care for the young. Such foster

mothers are typically prepared by mating with vasectomized males of the same species. The stage of the pseudopregnant foster mother is important for successful implantation, and it is species dependent. For mice, this stage is about 2-3 days pseudopregnant.

5           Offspring that are born to the foster mother may be screened initially for mosaic coat color where the coat color selection strategy (as described above, and in the appended examples) has been employed. In addition, or as an alternative, DNA from tail tissue of the offspring may be screened for the presence of the knockout construct using Southern blots and/or PCR as described above. Offspring that appear  
10   to be mosaics may then be crossed to each other, if they are believed to carry the knockout construct in their germ line, in order to generate homozygous knockout animals. Homozygotes may be identified by Southern blotting of equivalent amounts of genomic DNA from mice that are the product of this cross, as well as mice that are known heterozygotes and wild type mice.

15           Other means of identifying and characterizing the knockout offspring are available. For example, Northern blots can be used to probe the mRNA for the presence or absence of transcripts encoding either the gene knocked out, the marker gene, or both. In addition, Western blots can be used to assess the level of expression of the *RAFTK* gene knocked out in various tissues of the offspring by probing the  
20   Western blot with an antibody against the particular *RAFTK* protein, or an antibody against the marker gene product, where this gene is expressed. Finally, in situ analysis (such as fixing the cells and labeling with antibody) and/or FACS (fluorescence activated cell sorting) analysis of various cells from the offspring can be conducted using suitable antibodies to look for the presence or absence of the knockout construct  
25   gene product.

          Yet other methods of making knock-out or disruption transgenic animals are also generally known. See, for example, *Manipulating the Mouse Embryo* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Recombinase dependent knockouts can also be generated, e.g. by homologous recombination to  
30   insert target sequences, such that tissue specific and/or temporal control of inactivation of a *RAFTK*-gene can be controlled by recombinase sequences (described infra).

          Animals containing more than one knockout construct and/or more than one transgene expression construct are prepared in any of several ways. The preferred  
35   manner of preparation is to generate a series of mammals, each containing one of the desired transgenic phenotypes. Such animals are bred together through a series of crosses, backcrosses and selections, to ultimately generate a single animal containing

all desired knockout constructs and/or expression constructs, where the animal is otherwise congenic (genetically identical) to the wild type except for the presence of the knockout construct(s) and/or transgene(s).

The present invention is further illustrated by the following examples which should not be construed as limiting in any way. The contents of all cited references, including literature references, issued patents, published patent applications as cited throughout this application are hereby expressly incorporated by reference. The practice of the present invention will employ, unless otherwise indicated, conventional techniques of cell biology, cell culture, molecular biology, transgenic biology, microbiology, recombinant DNA, and immunology, which are within the skill of the art. Such techniques are explained fully in the literature. See, for example, *Molecular Cloning A Laboratory Manual*, 2nd Ed., ed. by Sambrook, Fritsch and Maniatis (Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning. Volumes I and II* (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Mullis et al. U.S. Patent No: 4,683,195; *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds. 1984); *Transcription And Translation* (B. D. Hames & S. J. Higgins eds. 1984); *Culture Of Animal Cells* (R. I. Freshney, Alan R. Liss, Inc., 1987); *Immobilized Cells And Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide To Molecular Cloning* (1984); the treatise, *Methods In Enzymology* (Academic Press, Inc., N.Y.); *Gene Transfer Vectors For Mammalian Cells* (J. H. Miller and M. P. Calos eds., 1987, Cold Spring Harbor Laboratory); *Methods In Enzymology*, Vols. 154 and 155 (Wu et al. eds.), *Immunochemical Methods In Cell And Molecular Biology* (Mayer and Walker, eds., Academic Press, London, 1987); *Handbook Of Experimental Immunology, Volumes I-IV* (D. M. Weir and C. C. Blackwell, eds., 1986); *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986).

**EXAMPLE 1: Identification and Characterization of a Novel Related Adhesion Focal Tyrosine Kinase (RAFTK) from Megakaryocytes and Brain**

30

A cDNA encoding a novel human intracytoplasmic tyrosine kinase, termed *RAFTK* (for a related adhesion focal tyrosine kinase) was isolated. The murine homolog of the human *RAFTK* cDNA was also cloned and characterized. Comparison of the deduced amino acid sequences of human and murine *RAFTK* cDNAs revealed 95% homology, indicating that *RAFTK* is highly conserved between these species. The *RAFTK* cDNA clone, encoding a polypeptide of 1009 amino acids, has closest homology (48% identity, 65% similarity) to the focal adhesion kinase

35

(pp125<sup>FAK</sup>). Comparison of the deduced amino acid sequences also indicates that *RAFTK* like pp125<sup>FAK</sup> lacks a transmembrane region, myristylation sites and SH2 and SH3 domains. In addition, like pp125<sup>FAK</sup>, *RAFTK* contains a kinase domain flanked by large N-terminal (426 residues) and C-terminal (331 residues) domains, and the C-terminal region contains a predicted proline-rich stretch of residues. In fetal tissues, *RAFTK* expression was abundant in brain and low levels were observed in lung and liver. In adult tissues, it was less restricted indicating that *RAFTK* expression is developmentally upregulated. Expression of *RAFTK* was also observed in human CD34<sup>+</sup> marrow cells, primary bone marrow megakaryocytes, platelets and various areas of brain. The human *RAFTK* gene was assigned to human Chromosome 8 using genomic DNAs from human/rodent somatic cell hybrid lines. The mouse *RAFTK* gene was mapped to Chromosome 14 closely linked to gonadotropin releasing hormone. Using specific antibodies for *RAFTK*, an approximately 123 Kd protein from the human CMK megakaryocyte cell line was immunoprecipitated. Treatment of the CMK megakaryocytic cells with thrombin caused a rapid induction of tyrosine phosphorylation of *RAFTK* protein. The structural features of *RAFTK* suggest that it is a member of the focal adhesion kinase gene family and may participate in signal transduction in human megakaryocytes and brain as well as other cell types.

The predicted amino acid sequence of the *RAFTK* protein shares consensus motifs in the central catalytic domain common to protein tyrosine kinases. The *RAFTK* cDNA, encoding a polypeptide of 1009 amino acids, has the closest homology (48% identity, 65% similarity) to FAK. Analysis of their deduced amino acid sequences also indicates that *RAFTK*, like FAK, lacks a transmembrane region, myristylation sites, and SH2 and SH3 domains. In addition, like FAK, the *RAFTK* C-terminal domain contains a predicted proline-rich stretch of residues. *RAFTK* was reported to be highly expressed in the central nervous system (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10) and involved in Ca<sup>2+</sup>-induced regulation of ion channel and MAP kinase functions in PC-12 cells (Lev, S. et al. (1995) *Nature* 376, 737).

*RAFTK* expression is abundant in primary bone marrow megakaryocytes and their progeny, blood platelets (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10). To address the role of *RAFTK* in signal transduction pathways in megakaryocytes, experiments were performed using the model CMK megakaryocytic cell line (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485). The c-kit receptor and its cognate ligand SCF were investigated since they play a critical role in the adhesion, migration, motility, proliferation and maturation of a number of hematopoietic cells, including megakaryocytes and platelets (See, e.g., Dastyh, J. and Metcalfe, D. D. (1994) *J.*

*Immunol.* 152, 213-219; Kinashi, T. et al. (1995) *Blood* 86, 2086-2090; Scott, G. et al. (1994) *Pigment Cell Res.* 7, 44-51). Since SCF effects appear to be modulated by PKC (Blume-Jensen, P. et al. (1995) *Journal of Biological Chemistry* 270, 14192-14200), the ability of PKC to mediate the effects of SCF and  $\text{Ca}^{2+}$  on *RAFTK* phosphorylation was investigated. In this study, SCF and PMA induced the tyrosine phosphorylation of *RAFTK* through PKC. In addition, *RAFTK* was associated with the cytoskeletal protein paxillin in megakaryocytes, and this association appeared critical for *RAFTK* phosphorylation.

The following materials and methods were used to clone and characterize *RAFTK*

### Materials

Chemical reagents were purchased from Sigma (St. Louis, MO). Restriction endonucleases, modifying enzymes, and terminal deoxynucleotidyl transferase were purchased from Pharmacia Biotech, Inc. (Piscataway, NJ) and New England Biolabs (Beverly, MA). The primers for polymerase chain reaction (PCR), RNA-PCR and sequencing were synthesized by an automated DNA synthesizer (Applied Biosystems, model 394). The PCR and RNA-PCR reagents were obtained from Perkin-Elmer Cetus (Norwalk, CT) and random primed labeling kits were obtained from Stratagene (La Jolla, CA). Manual and automated sequencing kits were obtained from USB (Cleveland, OH) and Pharmacia Biotech, Inc., respectively. Automated sequencing was performed using Pharmacia's Automated Laser Fluorescent Sequencer (ALF). Monoclonal antibody 2A7 against pp125<sup>FAK</sup> protein was kindly obtained from Dr. J. Thomas Parsons (Charlottesville, VA). Monoclonal antibody PY-20 directed against Tyr(P) was obtained from ICN (Costa Mesa, CA).

### Cells

Human marrow megakaryocytes were isolated by a method employing immunomagnetic beads using anti-human glycoprotein GpIIb monoclinal antibody, as previously described (Tanaka, H. et al. (1989) *Br. J. Haematol.* 73, 18-22; Avraham, H. et al. (1992) *Blood* 80, 1679-1684). CD34 bearing marrow progenitor cells were purified from heparinized bone marrow aspirates using immunomagnetic beads coated with an anti-CD34 monoclonal antibody as previously described (Avraham, H. et al. (1992) *Blood* 80, 1679-1684). The CMK cell line, provided by Dr. T. Sato and derived from a child with megakaryoblastic leukemia, has properties of cells of the megakaryocytic lineage (Sato, T. et al. (1987) *Exp. Hematol.* (N.Y.) 15,

495-502). The CMK cell line was cultured in RPMI 1640 medium with 10% fetal calf serum. Additional permanent human megakaryocytic cell lines were studied. DAMI cells were obtained from Dr. S. Greenberg, (Brigham and Women's Hospital, Boston, MA), Mo7e and erythroid-megakaryocytic HEL cells were obtained from Dr. L. Zon, (Children's Hospital, Boston, MA). Each cell line was cultured as previously described (Avraham, H. et al. (1992) *Blood* 80, 1679-1684; Avraham, H. et al. (1992) *Blood* 79, 365-371; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-75). Other permanent human cell lines such as Ramos (human B-cells) were obtained from the American Type Tissue Culture Collection and maintained in liquid culture according to the specifications in the catalog.

Human platelets were isolated by gel filtration from freshly drawn blood anticoagulated with 0.15 vol NIH formula A acid-citrate-dextrose solution supplemented with 1  $\mu$ M prostaglandin E<sub>1</sub> (PGE<sub>1</sub>) as previously described (Lipfert, L. et al. (1992) *J. Cell Biol.* 119, 905-912).

15

#### DNA amplification and cloning

Total RNA derived from CMK cells was prepared by a standard protocol of lysis in guanidinium isothiocyanate followed by cesium chloride gradient centrifugation (Maniatis, T. et al. (1992) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). Protein-tyrosine kinase sequences were amplified with degenerate oligonucleotide primers as previously described (Wilks, A.F. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1603-1607).

Briefly, total RNA (10  $\mu$ g) was used as a template for synthesis of complementary DNA (cDNA). The PTK3 oligonucleotide "SDVWSF/YG" (SEQ ID NO:5) 5'- (C/G)(T/A)(A/G)TC(A/C/G/T)ACCCA(A/C/G/T)(C/G)(T/A)(A/G)(T/A)A (A/C/G/T)CC - 3' (SEQ ID NO:6) was designed in our lab and was used as a primer. PCR was performed on one quarter of the cDNA synthesis reaction mixture (original volume-20  $\mu$ l), using PTK1 "DLAARN" (SEQ ID NO:7) 5'- CGACGA(T/C)CT(A/C/G/T)GC(A/C/G/T) (A/G)C(A/C/G/T)AA - 3'(SEQ ID NO:8) and PTK2 "WMAPE" (SEQ ID NO:9) 5' - GTACC(T/C)TC(G/C/A)GG(A/C/G/T)GCCATCCA - 3' (SEQ ID NO:10) oligonucleotides (50 pmol each) (Wilks, A.F. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1603-1607). The mixture was then subjected to PCR amplification using the Perkin-Elmer Cetus thermal cycler set for 30 cycles as follows: denature 95°C, 2 min; primer anneal 37°C, 1.5 min; primer extension 72°C, 2.30 min; 1 minute ramp times were used between these temperatures. PCR products of the amplified tyrosine kinase domains were purified from the agarose gel, digested with EcoRI and BamHI, ligated

into pUC19, and transformed into *Escherichia coli* DH5 $\alpha$ . Sequencing was carried out by the dideoxy chain termination method using version 2.0 sequenase kit (USB, Cleveland, OH). Sequences were compared with known sequences in GenBank and EMBL data bases using the Autosearch computer program. A novel clone was  
5 identified. This 160-base pair (bp) PCR product, designated JJ3, was radiolabeled using the "Prime It II" random priming protocol (Stratagene) and used as a probe to screen human cDNA libraries.

#### Isolation and characterization of cDNA clones

10 The human brain (hippocampus) cDNA library in  $\lambda$ -ZapII vector (randomized and oligo dT, cat #936205, Stratagene, CA) was screened ( $\sim 5 \times 10^5$  recombinants/screening) initially with the 160 bp PCR fragment (termed JJ3), and labeled with [ $\gamma^{32}\text{P}$ ] dCTP using random primed cDNA labeling. Hybridization to  
15 nylon filters (MSI) was performed in 50% formamide, 6 x SSC, 10 mM sodium phosphate, 5 x Denhardt's solution, 0.1% sodium dodecyl sulfate (SDS), and 1 mg/ml Herring sperm DNA (Boehringer Mannheim, Germany) at 43°C overnight. The filters were washed at room temperature in 2 x SSC, 1% SDS, and then in 0.2 x SSC, 0.1% SDS at 63°C three times for 30 min., UV crosslinked (Stratagene Stratalinker), and exposed to Kodak X-OMAT AR film (Eastman Kodak). Twelve clones were  
20 isolated and processed. Plasmid DNA was prepared using Exassist Helper Phage and the SolR System according to the manufacturer's instructions (Stratagene). Of these twelve clones, two were sequenced on both strands. A human CMK-PMA cDNA library oligo dT (Avraham, H. et al. (1992) *Blood* 79, 365-371) ( $\sim 3 \times 10^5$  recombinants/screening) in  $\lambda$ -gt10 vector was screened with the  $^{32}\text{P}$ -labeled JJ3  
25 fragment. Four clones were isolated and the recombinant DNAs of 2 positive phages were digested with EcoRI, and the cDNA insert was subcloned into pBSK (Stratagene) and thereafter sequenced.

A 340 bp probe was prepared from the 5'- end of one of the CMK cDNA clones (termed 2-1) and used to screen the human brain (hippocampus) cDNA library.  
30 Twelve clones were isolated and two clones were sequenced on both strands. In addition, a 248 bp probe was prepared from the 5'- end of one of the clones (termed 4C) and the human hippocampus cDNA library was rescreened. Twelve clones were identified and isolated and of these, 1 clone (termed 3B) was sequenced on both strands.

35 The mouse brain cDNA library (cat # ML1042b, Clontech, Palo Alto, CA) in  $\lambda$ -gt11 vector was screened ( $\sim 5 \times 10^5$  recombinants/screening) using 381 bp 5'- Kpn I fragment or 764 bp ApaI -3'- fragment of human *RAFTK* cDNA as a probe and the

filters were hybridized and washed under high stringency conditions. Six clones were isolated. The DNA was isolated as previously described (Maniatis, T. et al. (1992) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY) and subcloned into pBSK and thereafter sequenced.

- 5           Nucleotide sequences were determined by the Automated Laser Fluorescent (ALF) DNA sequencer (Pharmacia Biotech, Inc.) using Autoread (Pharmacia) and by manual sequencing using sequenase kit (USB).

### **Chromosomal localization of the human *RAFTK* gene**

- 10           Genomic DNAs from the NIGMS Hybrid Mapping Panels #1 and #2 were obtained from the NIGMS Genetic Mutant Cell Repository (Coriel Cell Institute for Medical Research, Camden, NJ). In addition, both mapping panels included DNA samples isolated from human and rodent parental cell lines (mouse and Chinese hamster). Approximately 5 µg of DNA from human, hamster, and mouse genomic  
15           DNAs were digested with BamHI, HindIII and PstI to find a suitable restriction fragment length polymorphism (RFLP) or unique genomic fragment for use in mapping. Subsequently, genomic DNAs from each panel were cut with BamHI. Southern blots were probed with a human 1.4 kb *RAFTK* cDNA and hybridizations were carried out as previously described (Rowe, L.B. et al. (1994) *Mamm. Genome* 5,  
20           253-274; White, R.A. et al. (1992) *Nature Genet.* 2, 80-83). Hybrids were scored for the appropriate human-specific restriction endonuclease fragment on the autoradiographs. The results were compared with the chromosome contents of the hybrid cell lines and the concordance between restriction fragments and specific chromosome content was used to establish the localization of human *RAFTK*.

25

### **Backcross mapping of the mouse *RAFTK* gene**

- Genomic DNAs from C57BL/6J, *Mus spretus* and a (*M. spretus* x C57BL/6J) *M. spretus* BSS type backcross DNA panel were obtained from The Jackson Laboratory (Bar Harbor, Maine) (Rowe, L.B. et al. (1994) *Mamm. Genome* 5, 253-  
30           274). Southern blots and hybridizations were performed as previously described (White, R.A. et al. (1992) *Nature Genet.* 2, 80-83). Approximately 5 µg of genomic DNAs of C57BL/6J and *Mus spretus* were digested with 29 different restriction enzymes to identify a potential RFLP genetic marker. The Southern blots were probed with a 1.4 kb human *RAFTK* cDNA fragment labeled with <sup>32</sup>P using a  
35           Decaprime II Kit (Ambion, Inc., Austin, TX). Digestion of the backcross DNA panel with BamHI, Southern blotting and hybridizations were carried out as previously described (White, R.A. et al. (1992) *Nature Genet.* 2, 80-83).



### Recombinant inbred (RI) line mapping of the mouse *RAFTK* gene

*RAFTK* and *Gnrh* co-segregated in BXD RI lines and mapped to chromosome 14. Genomic DNAs isolated from the progenitors of BXD RI lines (C57BL/6J and DBA/2J) were digested with 29 different restriction enzymes to identify a RFLP genetic marker for mapping. Subsequently, genomic DNAs isolated from the BXD RI lines were digested with *SacI*. Conditions for Southern blots and hybridizations were the same as previously described (White, R.A. et al. (1992) *Nature Genet.* 2, 80-83) and the 1.4 kb human *RAFTK* cDNA was used as a probe. Data were compared with strain distribution patterns (SPDs) recorded in GBASE (1993) (Yang-Feng, T.L. et al. (1986) *Somatic Cell. Mol. Genet.* 12, 95-100).

### Northern blot analysis

Total RNA was prepared by a standard protocol of lysis in guanidinium isothiocyanate followed by cesium chloride gradient centrifugation (Maniatis, T. et al. (1992) *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY). The human adult and fetal tissue Northern blots, the brain regions and the human tissue II blots were obtained from Clontech (Palo Alto, CA). Hybridization was carried out according to the manufacturer's instructions. Each RNA blot was probed with a 146 bp, 3'- gene-specific *RAFTK* cDNA radiolabeled to a high specific activity ( $10^8$ - $10^9$  cpm/ $\mu$ g) with [ $\gamma$ - $^{32}$ P]dCTP. The level of expression for each mRNA was also determined densitometrically (EC Apparatus Corp. Densitometer; St. Petersburg, FL). The radioactivity associated with each band was also assayed with a Betascope 603 blot analyzer (Betagen, Mountain View, CA). The same blot was assessed for the presence of the actin or GAPDH specific probes; Actin and GAPDH were used as controls to assure uniform loading.

### PCR blots

cDNA was prepared from platelets ( $10 \times 10^8$ ), CD34<sup>+</sup> marrow cells ( $10^6$  cells) and bone marrow megakaryocytes ( $10^6$  cells) and amplified by PCR using specific *RAFTK* primers as previously described (Bennett, B.D. et al. (1994) *J. Biol. Chem.* 269, 1068-1074). The sequence of the *RAFTK* upstream primer was 5'-CGGGCCGTGCTGGAGCTCAA - 3' (SEQ ID NO:11)(position 2958 - 2977). The nucleotide sequence of the *RAFTK* downstream primer was 5'-GTCCGTGAAGATGACGGCAA - 3' (SEQ ID NO:12) (position 3084 - 3103). The sequence of the FAK upstream primer was 5'- AAAGCTGTCATCGAGATGTCC -3' (SEQ ID NO:13) (position 2292-2312). The nucleotide sequence of the downstream

primer was 5'- TCGGTGGGTGCTGGCTGGTAGG - 3' (SEQ ID NO:14) (position 2417-2438)(Andre, E., and Becker-Andre, M. (1993) *Biochem. Biophys. Res. Commun.* 190, 140-147). The sequence of the actin upstream primer was 5'- ATCTGGCACCACACCTTCTACAATGAGCTGCG - 3' (SEQ ID NO:15). The  
5 nucleotide sequence of the downstream primer was 5'- CGTCATACTCCTGCTTGCTGATCCACATCTGC -3' (SEQ ID NO:16) (Clontech, Palo Alto, CA). The PCR products were electrophoresed on a 1.5% agarose gel, denatured, neutralized, transferred to filters, and vacuum blotted. The probes used were the *RAFTK*, FAK and actin gene-specific probes, which were labeled by random  
10 priming as described above. Prehybridization and hybridization were carried out as previously described (Bennett, B.D. et al. (1994) *J. Biol. Chem.* 269, 1068-1074).

### Protein analysis

Metabolic labeling, immunoprecipitation, and Western blot analysis were  
15 performed in CMK cells as previously described (Laemmli, U.K. (1970) *Nature* 227, 680-685; Yarden, Y et al. (1987) *EMBO J.* 6, 3341-3351; Konopka, J. B., and Witte, O.N. (1985) *Mol. Cell Biol.* 5, 3116-3123; Konopka, J. B. et al. (1984) *J. Virol.* 51, 223-232). For immunoblot analysis, total cell lysates of CMK cells untreated or stimulated with  $\alpha$ -thrombin (1 U/ml or 2 U/ml as indicated)(ChromoLog Corp.,  
20 Havertown, PA) for 5 min were prepared as previously described (Yarden, Y et al. (1987) *EMBO J.* 6, 3341-3351). Relative protein concentrations were determined with a colorimetric assay kit (Bio-Rad Laboratories, Inc., Hercules, CA) with bovine serum albumin as the standard. A portion of lysate containing approximately 0.05 mg of protein was mixed with an equal volume of 2 x SDS sample buffer containing  $\beta$ -  
25 mercaptoethanol, boiled for 5 min., fractionated on 8% polyacrylamide-SDS gels (Laemmli, U.K. (1970) *Nature* 227, 680-685) and transferred to Immobilon polyvinylidene difluoride (Millipore Corp., Bedford, MA) filters. Protein blots were treated with specific *RAFTK* antibodies (R-4250) (see below). Primary binding of the *RAFTK* antibodies (see below) was detected using anti-IgG second antibodies  
30 conjugated to horseradish peroxidase and subsequent chemiluminescence development using the ECL Western blotting system (Amersham Life Sciences, Arlington Heights, IL).

For metabolic labeling,  $10^6$  cells were labeled with 100  $\mu$ Ci of [ $^{35}$ S] methionine in 1 ml of Dulbecco's modified Eagle's medium minus methionine  
35 (Amersham Life Sciences) for 16 h. Immunoprecipitation of *RAFTK* protein from labeled cells with *RAFTK* antiserum or with normal rabbit serum (NRS) was performed as previously described (Bennett, B.D. et al. (1994) *J. Biol. Chem.* 269,

1068-1074; Yarden, Y et al. (1987) *EMBO J.* 6, 3341-3351). For immunoprecipitation of Tyr (P) proteins, cold soluble extracts were first incubated with *RAFTK* antibodies (R-4250) overnight at 4°C. The extracts were then incubated with protein-G-Sepharose beads precoupled to goat anti-rabbit IgG for 1.5 h at 4°C.

- 5 Proteins were eluted from the beads by heating the samples at 100°C for 5 min in SDS-polyacrylamide gel electrophoresis buffer. Proteins were separated by SDS-PAGE, transferred and immunoblotted with PY-20 (diluted 1:5000). The immunoreactive bands were visualized using the ECL system.

## 10 Antibodies

- Anti-*RAFTK* antiserum was obtained from New Zealand white rabbits immunized with a bacterially expressed fusion protein consisting of the GST-C-terminal (681-1009 amino acid residues) of human *RAFTK* cDNA subcloned into the pGEX-2T expression vector. The sera were titrated against the GST-*RAFTK* C-terminus fusion protein by ELISA (Dymecki, S. M. et al. (1992) *J. Biol. Chem.* 267, 4815-4823; Bennett, B.D. et al. (1991) *J. Biol. Chem.* 266, 23060-23067) and the serum (R-4250) exhibiting the highest titer (1:256,000) was used in subsequent experiments.

## 20 Isolation and characterization of *RAFTK* cDNAs

- To identify tyrosine kinases in human megakaryocytes, PCR primers based on conserved sequences of PTKs were used (Wilks, A.F. (1989) *Proc. Natl. Acad. Sci. USA* 86, 1603-1607). RNA from the human megakaryocytic CMK cell line was used as a template to synthesize CMK cDNA. The cDNA was amplified by using the PTK primers. Fragments of the expected size (~160 bp) were isolated and subcloned for sequence analysis. One clone that appeared to represent a novel tyrosine kinase (termed JJ3) was used as a probe to screen the human hippocampus cDNA library. A partial cDNA clone (termed S2-3) containing an ~2.0 kb insert was isolated. A homology analysis of this clone to human pp125<sup>FAK</sup> was performed and regions were chosen to design specific primers to generate an *RAFTK* gene-specific probe. The JJ3 fragment was used to screen the human hippocampus cDNA library to obtain overlapping cDNAs. The 5'- end of each of these clones was in turn used as probes to obtain the full-length *RAFTK* cDNA. Eight different overlapping sequences were obtained of the coding region of *RAFTK*. Figure 1 is a schematic representation along with a restriction map of the sequence showing the pattern of overlapping cDNAs. The 3.6 kb length of the *RAFTK* cDNA contains an open reading frame with the first in frame ATG codon located at nucleotide 294-296, followed by a stop codon at

position 3260-3262. This open reading frame encodes a predicted protein of 1009 amino acid residues with a calculated molecular weight of ~123 Kd and has been given the name *RAFTK* (for a related adhesion focal tyrosine kinase). Analysis of the hydrophobicity of the predicted protein revealed lack of a transmembrane region and no recognizable sites for acylation. The kinase domain is flanked by large N-terminal (426 residues) and C-terminal (331 residues) domains. Comparison of the nucleotide sequence and the deduced amino acid sequence of the encoded protein with the National Biomedical Research Foundation and GenBank data bases revealed that this cDNA encoded a tyrosine kinase related to the pp125<sup>FAK</sup>. The predicted amino acid sequence of pp120<sup>RAFTK</sup> contains the structural motifs common to all protein kinases, including the putative ATP-binding site (432-437a.a, Gly<sup>432</sup> - Xaa - Gly<sup>434</sup> - Xaa - Xaa - Gly<sup>437</sup>), and three residues that are predicted to interact with the  $\gamma$ -phosphate group of the bound ATP molecule (in positions 402a.a, 529a.a and 655a.a). In addition, *RAFTK* contains two peptide sequences that are highly conserved among PTKs (Asp<sup>549</sup> - Ile<sup>550</sup> - Ala<sup>551</sup> - Val<sup>552</sup> - Arg<sup>553</sup> - Asn<sup>554</sup> and Pro<sup>588</sup> - Ile<sup>589</sup> - Lys<sup>590</sup> - Trp<sup>591</sup> - Met<sup>592</sup>). Interestingly, like chicken pp125<sup>FAK</sup>, the C-terminal region of *RAFTK* contains a proline-rich stretch (residues 690-767) where the proline content exceeds 20%. A unique domain is found at the NH<sub>2</sub> terminus of *RAFTK* (amino acids 1-39) (Figure 3). This region is the most divergent among various PTKs and may be involved in cellular localization and/or interaction with other cellular proteins. Like pp125<sup>FAK</sup>, *RAFTK* does not contain SH2 or SH3 domains. The kinase domain (amino acid 427 - 679) of *RAFTK* shares 60% identical homology with the mouse pp125<sup>FAK</sup>, 54% with human pp125<sup>FAK</sup>, and 36% with src (Figure 2). The kinase domain consists primarily of the catalytic domain including the putative ATP binding site (amino acids 432 - 437). *RAFTK* shares 42% homology in the NH<sub>2</sub> domain and about 39% in the C-terminal domain with mouse pp125<sup>FAK</sup>. The overall amino acid homology of *RAFTK* is 48% identity (65% similarity) with mouse pp125<sup>FAK</sup>.

### 30 Molecular cloning of the full-length murine *RAFTK* cDNA

Southern blot analysis of human and mouse genomic DNA digested with EcoRI, HindIII, BamHI, XbaI, PstI and probed under conditions of high stringency with 3'- fragment of *RAFTK* cDNA from 1595 - 2974 bp (1.4 kb) as a probe, revealed a single band in each lane, indicating that the human *RAFTK* gene and the mouse *RAFTK* gene are highly homologous and are single genes. Therefore, a random and oligo (dT)-primed mouse adult brain cDNA library was screened under conditions of high stringency for the full-length mouse cDNA of *RAFTK* using the 5'- fragment and

3'- fragment of human *RAFTK* cDNA as probes. Four clones were isolated and two of these clones were sequenced in both directions and additional clones were partially sequenced. Sequence analysis of these clones revealed identical sequences. The 4.5 kb full length cDNA has an open reading frame of 1009 amino acid residues and possesses 95.6% identical homology with the human *RAFTK* gene.

### Chromosomal localization of human *RAFTK* gene

Hamster, human and mouse DNAs were digested with BamHI, HindIII and PstI to identify a specific RFLP pattern for the *RAFTK* gene in each species. Southern blots were probed with a human 1.4 kb *RAFTK* cDNA. Unique BamHI 16.5 kb and 14.5 kb fragments for *RAFTK* were identified in human DNA from the parental cell lines used to prepare human/rodent cell hybrids. DNAs from the parental and the somatic hybrid cell lines in mapping panel #2 were digested with BamHI, Southern blotted and probed. Analysis indicated that the human-specific BamHI pattern was observed in cell line #8 which contains human Chromosome 8 (Figure 4). A fainter signal was also observed for the human-specific BamHI pattern in hybrid cell line #20 (Fig. 4) which, although it contained an intact human Chromosome #20, also carried a gene from human Chromosome #8 (NEFL, neurofilament light polypeptide, 8p21) as determined by Southern blot hybridization (Coriel Cell Institute for Medical Research, Camden, NJ). All other hybrid cell lines were negative for the human-specific BamHI RFLP. Additionally, when the human 1.4 kb *RAFTK* cDNA was used to probe Coriel Panel #1, the human-specific fragment was detected in all hybrids containing greater than 4% of human Chromosome 8 and was absent in every hybrid that lacked Chromosome 8.

Southern blots of C57BL/6J and *Mus spretus* DNAs were digested with 29 different restriction enzymes and probed with a human *RAFTK* 1.4 kb cDNA. A BamHI restriction fragment length polymorphism (RFLP) was detected. The alleles for this BamHI RFLP consist of 8.6 kb and 5.2 kb genomic DNA bands, characteristic of C57BL/6J, and 15.5 kb and 6.7 bands which are found in *Mus spretus*. These alleles were characterized in 87 DNAs from the C57BL/6J X *Mus spretus* backcross panel. Results of the haplotype analysis from this mapping data indicate that the *RAFTK* gene co-localizes with D14Bir10 (DNA segment-Birkenmeier 10) and is linked to Nfl (neurofilament, light polypeptide) on mouse Chromosome 14 (Figure 5). The *RAFTK* locus mapped between Xmv19 (xenotropic-MCF leukemia virus-19) and Nfl and the calculated map distances for these loci are: Xmv19,  $7.1 \pm 5.3$  cM, *RAFTK*,  $3.5 \pm 2.0$  cM, Nfl.

The position of *RAFTK* on mouse Chromosome 14 was confirmed by determining the segregation of a *SacI* RFLP for *RAFTK* DNAs from BXD recombinant inbred (RI) lines. The *SacI* RFLP for *RAFTK* was indicated by the presence of a 16.5 kb genomic DNA band in C57BL/6J or a 6.2 kb fragment in DBA/2J. These alleles were characterized for 26 DNAs from the BXD RI line. The strain distribution patterns of *RAFTK* and the locus coding for gonadotropin releasing hormone, *Gnrh* (Hearne, C.M. et al. (1991) *Mamm. Genome* 1, 273-282), indicate close linkage between these two loci on Chromosome 14. Perfect concordance was observed with the BXD strain distribution pattern for the *Gnrh* locus, indicating linkage of less than one map unit distance from *RAFTK* *Gnrh* (Silver, J. (1985) *J. Hered.* 76, 436-440). This mapping data places *RAFTK* distal to *Nfl* and is a contradiction to the backcross data. However, backcross data are not as accurate as RI data since backcross mice were derived from interspecies cross.

#### 15 **Expression of *RAFTK* in tissues and cell lines**

A specific *RAFTK* probe was designed (nucleotide 2958 bp - 3103 bp). This sequence is present in *RAFTK* and not in human pp125<sup>FAK</sup>. This probe was used for hybridization of all Northern blots described here.

Northern blot analysis of RNA from human fetal heart, brain, lung, liver and kidney revealed a weak single major species of mRNA of 4.5 kb in brain and it appears to be expressed at low levels in the lung and liver. Expression in human adult tissues was assessed by hybridization of the cDNA probe to a Northern blot of poly (A<sup>+</sup>) RNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. While heart and skeletal muscle RNA samples were negative for *RAFTK*, a single mRNA was observed in all other tissues with the highest levels expressed in brain. To further characterize the distribution of *RAFTK* expression in other human tissues, Northern blot analysis of spleen, thymus, prostate, testes, ovary, intestine, colon and peripheral blood leukocytes revealed high expression of *RAFTK* in thymus, spleen and peripheral blood leukocytes. Northern blot analysis of different human brain regions (amygdala, caudate nucleus, corpus callosum, hippocampus, hypothalamus, substantia nigra, subthalamic nucleus and thalamus) revealed that the highest expression of *RAFTK* was in amygdala and hippocampus. Lower expression was observed in the other brain regions, with the exception of corpus callosum and substantia nigra where there was no detectable signal. These results demonstrate that brain has abundant expression of *RAFTK*, especially in the amygdala and hippocampus.

Expression of *RAFTK* was observed in several megakaryocytic cell lines such as CMK, Mo7e, HEL and DAMI cells. In addition, expression of *RAFTK* was detected in Ramos, FHS and HeLa cells but low level of expression was detected in Jurkat, Hep 3B and CCL 75 cells. Using PCR techniques, expression of *RAFTK* was also found in primary bone marrow megakaryocytes, blood platelets, and in marrow CD34<sup>+</sup> progenitor cells. The level of expression of *RAFTK* mRNA is similar to FAK in CD34<sup>+</sup> cells, and is higher than FAK in bone marrow megakaryocytes. In platelets, the level of expression of *RAFTK* mRNA is lower than FAK, as observed by PCR under the same experimental conditions. *RAFTK* mRNA expression in bone marrow megakaryocytes is higher than that in CD34<sup>+</sup> cells. Taken together, these results demonstrate that *RAFTK* is abundantly expressed in brain and hematopoietic cells. The restricted expression observed in fetal versus adult tissues indicates its expression is upregulated during development.

#### 15 **Generation of specific antibodies for *RAFTK* and detection of *RAFTK* protein**

The fusion protein GST-C-terminus of *RAFTK* (residues 681-1009) was chosen for rabbit immunizations in order to obtain specific antibodies for *RAFTK* protein. These polyclonal antibodies (R-4250) do not cross react with pp125<sup>FAK</sup>. The monoclonal antibody 2A7 against FAK does not cross react with the C-terminal GST-*RAFTK* fusion protein, indicating that *RAFTK* might be antigenically different from FAK. Furthermore, FAK immunoprecipitated by the monoclonal antibody 2A7 from megakaryocytes was not recognized by polyclonal antiserum 4250. Similarly, *RAFTK* immunoprecipitated by antiserum 4250 also was not recognized by the monoclonal antibody 2A7. Taken together, these data demonstrate that FAK and *RAFTK* are distinguishable antigenically while being related members of the FAK family.

The specificity of this antiserum was examined by immunoprecipitation. The CMK cell line was metabolically labeled with [<sup>35</sup>S] methionine, and extracts were immunoprecipitated with anti-*RAFTK* antiserum. A major protein species of ~123 Kd was detected in CMK cells. A similar species was observed in other human megakaryocytic cell lines such as DAMI. This band was not observed when normal rabbit serum or pre-immune rabbit serum was used for immunoprecipitation. Incubation of R4250 with 1 µg or 10 µg of the C-terminus of GST-*RAFTK* fusion protein abolished the appearance of ~123 Kd, while incubation with 10 µg of the fusion protein GST-MATK-SH2 domain did not have any effects. These results demonstrate that polyclonal antibodies R-4250 are specifically recognizing *RAFTK* protein of ~123Kd size. Furthermore, thrombin (1 unit/ml) stimulated a rapid increase

in the amount of *RAFTK* protein immunoreactivity in anti-Tyr(P) immunoprecipitates. These results demonstrate that *RAFTK* is a protein tyrosine kinase, and that thrombin can induce its tyrosine phosphorylation.

The method of PCR cloning has been successfully employed by many laboratories to identify novel members of the PTK family. Using this strategy, a novel intracytoplasmic tyrosine kinase in human megakaryocytic cells has been identified, termed *RAFTK*. Sequence analysis of *RAFTK* revealed ~48% identity (65% similarity) to pp125<sup>FAK</sup> suggesting that *RAFTK* belongs to this subfamily of cytoplasmic tyrosine kinases. *RAFTK* does not appear to be the recently described FAKB protein (Kanner, S.B. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91, 10484-10487), also related to pp125<sup>FAK</sup>, since the specific amino acid sequence used to make antisera which recognized the FAKB protein is missing in the predicted amino acid sequence of *RAFTK* protein. Furthermore, unlike FAKB, *RAFTK* protein did not form stable complexes with the TCR/CD3 linked tyrosine kinase ZAP 70 in T-cells indicating that *RAFTK* and FAKB are different proteins.

The chicken, human and mouse focal adhesion kinases have been recently implicated as playing key roles in signal transduction pathways associated with extracellular adhesion molecules and with receptors for neuropeptide growth factors (Schaller, M.D., and Parsons, J.T. (1993) *Trends Cell Biol.* 3, 258-262; Zachary, I., and Rozengurt, E. (1992) *Cell* 71, 891-894; Leeb-Lundberg, L.M., and Song, X.-H. (1991) *J. Biol. Chem.* 266, 7746-7749; Zachary, I. et al. (1992) *J. Biol. Chem.* 267, 19031-19034). Thus, based on its homology to pp125<sup>FAK</sup>, one would expect *RAFTK* to participate in signalling pathways as well. The deduced 1009 amino acid sequence of *RAFTK* (with calculated molecular mass of 120 Kd) contains a kinase domain and lacks a transmembrane region, myristylation sites, and SH2 and SH3 domains. In order to identify conserved regions within *RAFTK* between species that may have important functions, the murine homolog of the human *RAFTK* cDNA was cloned. The sequence identity between the human and murine *RAFTK* cDNAs is 90% at the nucleotide level and 95.6% at the predicted amino acid level. In the kinase domain, 98.5% of the amino acids are identical. Therefore, the *RAFTK* gene is highly conserved in human and rodent, again suggesting an important role in cell signalling functions. The *RAFTK* has an insertion of an additional 4 amino acids between 76-81 (G<sup>76</sup>R<sup>77</sup>I<sup>78</sup>G<sup>79</sup>) compared to chicken, murine, and human pp125<sup>FAK</sup> sequences (Schaller, M.D. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196; Schaller, M.D., and Parsons, J.T. (1993) *Trends Cell Biol.* 3, 258-262; Clark, E.A., and Brugge, J.S. (1995) *Science* 268, 233-239). Amino acids corresponding to positions 292 - 320 of human pp125<sup>FAK</sup> and amino acids corresponding to position 850 - 864 and 901 -



926 of chicken pp125<sup>FAK</sup> are absent in the predicted *RAFTK* protein. Interestingly, like chicken pp125<sup>FAK</sup>, the C-terminus region of human *RAFTK* and mouse *RAFTK* contains a proline-rich stretch (residues 690-767). It has been shown that proteins containing proline-rich peptide motifs (such as Shc, p62 and ribonucleoprotein K) could serve as SH3 domain ligands, and that the binding of these proteins to the Src SH3 domain was inhibited with a proline-rich peptide ligand (Weng, Z. et al. (1994) *Mol. Cell. Biol.* 14, 4509-4521). Furthermore, the predicted *RAFTK* protein, like the pp125<sup>FAK</sup> protein, displays several unique features among the known tyrosine kinases. The primary sequence of *RAFTK* does not contain a signal peptide or a membrane-spanning region and the protein is therefore presumed to be located in the cytoplasm. *RAFTK* lacks SH2 and SH3 domains, which are structural elements involved in protein-protein interactions (Pawson, T., and Gish, G.D. (1992) *Cell* 71, 359-362; Konopka, J. B. et al. (1984) *J. Virol.* 51, 223-232; Waksman, G. et al. (1992) *Nature* 358, 646-653; Taylor, S.J., and Shalloway, D. (1993) *Current Opin. Genet. & Dev.* 3, 26-34; Yu, H. et al. (1992) *Science* 258, 1665-1668), and does not exhibit significant homology with any known PTK beyond pp125<sup>FAK</sup> outside of the catalytic domain. Lack of SH2 and SH3 domains suggests that other regions within *RAFTK* protein are important for protein interaction and targeting. In the case of the pp125<sup>FAK</sup> protein, it has been demonstrated by structural-functional analysis that 159 amino acids within the C-terminus are essential as a "Focal adhesion targeting" sequence (Hildebrand, J.D. et al. (1993) *J. Cell. Biol.* 123, 993-1005). The homology between *RAFTK* and pp125<sup>FAK</sup> within this region is 52%. The overall structure of *RAFTK* is characteristic of the pp125<sup>FAK</sup> gene, with the catalytic domain flanked by large N-terminal and C-terminal domains. It has recently been reported that deletions of the NH<sub>2</sub> - or the COOH-terminal non-catalytic domain of pp125<sup>FAK</sup> including Tyrosine<sup>397</sup> did not abolish the kinase activity of pp125<sup>FAK</sup> (Chan, P.-Y. et al. (1994) *J. Biol. Chem.* 269, 20567-20574). Moreover, there is conservation of several tyrosine residues between *RAFTK* and pp125<sup>FAK</sup> including Tyrosine<sup>397</sup> which has been shown to be the major site of tyrosine phosphorylation in pp125<sup>FAK</sup> protein (Schaller, M.D. et al. (1994) *Mol. Cell. Biol.* 14, 1680-1688).

*RAFTK* specific mRNA expression was observed in human fetal tissues most abundant in brain (predominantly in amygdala and hippocampus regions) and appeared to be developmentally upregulated as demonstrated in the pattern of adult tissue expression. Within the hematopoietic system, in addition to peripheral blood leukocytes, a high level of specific mRNA expression of *RAFTK* was detected in B-

cells and various megakaryocytic cell lines. By using PCR, the specific mRNA expression of *RAFTK* was also detected in primary bone marrow CD34<sup>+</sup> progenitor cells, primary bone marrow megakaryocytes and platelets.

*RAFTK* is phosphorylated after thrombin treatment of CMK cells. FAK  
5 protein was also found phosphorylated on tyrosine after thrombin or collagen treatment of platelets (Lipfert, L. et al. (1992) *J. Cell Biol.* 119, 905-912). There is considerable homology in the thrombin receptors and considerable signal similarities in transduction mechanisms between platelets and megakaryocytes (Vittet, D., and Chevillard, C. (1993) *Blood Coagulation & Fibrinolysis* 4, 759-768). Furthermore,  
10 bone marrow megakaryocytes in liquid culture stimulated with thrombin for 5 min revealed dramatic morphological changes reminiscent of those found in platelets, including shape change and organelle centralization that involved immature as well as mature cells (Cramer, E.M. et al. (1993) *Am. J. Path.* 143, 1498-1508). Megakaryocytes were also able to secrete alpha-granule proteins in the dilated  
15 cisternae of the demarcation membrane system (Cramer, E.M. et al. (1993) *Am. J. Path.* 143, 1498-1508).

The human *RAFTK* gene was found on Chromosome 8 using DNAs from the somatic cell hybrid lines. The signal observed in cell line #20 in mapping panel #2 suggested that a fragment of Chromosome 8 is in the Chromosome #20 cell line.  
20 Although cell line #20 contained the human NEFL gene, there was no evidence for Chromosome 20 or a fragment of Chromosome 20 in cell line #8 (Coriel Cell Institute for Medical Research, Camden, NJ). The localization of *RAFTK* to Chromosome 8 was confirmed using mapping panel #1. The human NEFL gene has been localized to Chromosome 8p21 (Hurst, J. et al. (1987) *Cytogenet. Cell Genet.* 45, 30-32). Nfl, the  
25 murine homolog of human NEFL, has been mapped to mouse Chromosome 14 and is within 3 cM of the Gnrh locus (GBASE). The close linkage of the mouse *RAFTK* gene to Nfl (whose homolog NEFL is on human Chromosome 8p21) suggested that the human *RAFTK* gene may be mapped to Chromosome 8 based on homology between human and mouse chromosomes (Hurst, J. et al. (1987) *Cytogenet. Cell*  
30 *Genet.* 45, 30-32). Therefore, the human *RAFTK* gene is localized to Chromosome 8p21. The mouse *RAFTK* gene has been mapped to Chromosome 14 using a (C57BL/6J) x *M. spretus* F<sub>1</sub> x *M. spretus* backcross. The position of mouse *RAFTK* was confirmed by RI line mapping using the BXD RI lines. The *RAFTK* gene was also shown to be closely linked to Gnrh whose human homolog (LHRH-luteinizing  
35 hormone releasing hormone) has been mapped to human Chromosome 8p21-11.2 (Yang-Feng, T.L. et al. (1986) *Somatic Cell. Mol. Genet.* 12, 95-100).

**EXAMPLE 2:           Activation of the Novel Protein Tyrosine Kinase, *RAFTK*,  
in Megakaryocytes Upon SCF and PMA Stimulation and  
Its Direct Association with Paxillin**

5           *RAFTK* appears to be a member of the Focal Adhesion Kinase (FAK) family, and is involved in  $\text{Ca}^{2+}$ -mediated signalling events in PC-12 cells. In this Example, the signalling pathways involving *RAFTK* in human megakaryocytic cells were characterized. Stem Cell Factor (SCF), which potentiates the growth of megakaryocytes and their progenitors, and Phorbol Myristate Acetate (PMA), which  
10 causes differentiation of megakaryocytic cell lines, induced the tyrosine phosphorylation of *RAFTK* through Protein Kinase C (PKC). The constitutive association of *RAFTK* with PKC- $\delta$  was observed, while the association of *RAFTK* with PKC- $\alpha$  was induced upon stimulation with SCF. In addition, the direct association of *RAFTK* with paxillin, a 68-Kd cytoskeleton protein, was demonstrated.  
15 Upon the activation of *RAFTK*, there was a sequential activation and phosphorylation of paxillin. Cytochalasin D, which disrupts the cytoskeleton, abolished the phosphorylation of *RAFTK* upon PMA and SCF stimulation.

          These results show that *RAFTK* is a down-stream signalling protein of PKC and that paxillin is a down-stream associated protein of *RAFTK*. Furthermore,  
20 *RAFTK* association with the cytoskeleton was critical for its phosphorylation. These observations show the manner in which *RAFTK* participates in megakaryocyte proliferation and differentiation.

          The following materials and methods were used to study activation of *RAFTK*:  
25

**Materials**

          Recombinant SCF/KL and polyclonal anti-c-kit antibodies were generously provided by Dr. Keith E. Langley and Dr. L. Bennett, Amgen Inc. (Thousand Oaks, CA). Monoclonal anti-phosphotyrosine antibody (PY-20) and monoclonal anti-  
30 paxillin were obtained from ICN (Costa Mesa, CA ); monoclonal antibodies anti-p85, anti-Shc, anti-Grb2, anti-FAK, anti-PKC- $\alpha$ , anti-PKC- $\beta$ , and anti-PKC- $\delta$  were obtained from Transduction Laboratories (Lexington, KY). Calphostin C, staurosporine, calcium ionophore A23187, EGTA and Phorbol 12-Myristate 13-Acetate (PMA) were obtained from Calbiochem (La Jolla, CA ). Electrophoresis  
35 reagents were obtained from Bio-Rad Laboratories (Hercules, CA). All other reagents were purchased from Sigma Co. (St. Louis, MO).

### CMK cells

The CMK cell line, provided by Dr. T. Sato (Chiba University, Japan), was maintained in RPMI 1640 with 10% fetal calf serum (FCS) as described previously (Sato, T. et al. (1989) *Br. J. Hematol.* 72, 184-190). The CMK cell line was derived from a child with megakaryoblastic leukemia and has properties of cells of the megakaryocytic lineage, including the surface expression of glycoproteins Ib and IIb/IIIa, synthesis of platelet factor 4, PDGF and von Willebrand factor. CMK cells can proliferate in response to cytokines and have been used by us and other investigators in studies of megakaryocyte growth and maturation. In addition, CMK cells also differentiate upon induction with PMA (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-79; Namciu, S. et al. (1994) *Oncogen* 9, 1407-1416). For such experiments, PMA was dissolved in dimethyl sulfoxide and stored at -20°C until use, when it was diluted in RPMI 1640 medium.

### Antibodies

Anti-*RAFTK* antiserum was obtained from New Zealand White rabbits immunized with a bacterially expressed fusion protein consisting of GST and the C-terminus (amino acids 681-1009) of human *RAFTK* cDNA subcloned into the pGEX-2T expression vector as described (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10). The sera were titrated against the GST-*RAFTK* C-terminus fusion protein by an enzyme-linked immunosorbent assay, and the serum (R-4250) exhibiting the highest titer (1:256,000) was used in subsequent experiments. In immunoblotting experiments anti-phosphotyrosine antibodies were used (PY20) or anti-*RAFTK* (1:1000) or antibodies for FAK-2A7 (1:1000).

### Cell stimulation, immunoprecipitation and immunoblotting

The CMK cells were starved overnight in RPMI-1640 with 0.5% FCS. Cells (106/ml) were stimulated for 0 to 30 min at RT with either SCF (100 to 500 ng/ml) or PMA (10-100 nM). The stimulation was terminated by adding ice-cold RPMI-1640 containing sodium vanadate followed by centrifugation. The cells were lysed in modified-RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 2 (g/ml of aprotinin, leupeptin and pepstatin, and 1 mM Na<sub>3</sub>VO<sub>4</sub>). Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by using a protein assay (Bio-Rad Laboratories) and were standardized to equal concentrations of protein prior to immunoprecipitation. Identical amounts of protein

from each sample were precleared by incubation with Protein G-Sepharose CL-4B (Sigma Co., St. Louis, MO) for 1 h at 4°C. After the removal of Protein G-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as described below for each experiment for 4 h or overnight at 4°C.

- 5 Immunoprecipitation of the antigen-antibody complex was accomplished by incubation for 1 h at 4°C with 40 µl of protein G-Sepharose as described (Huang, E. et al. (1990) *Cell* 63, 225-233). Normal rabbit serum was used as a control in immunoprecipitations. Bound proteins were solubilized in 20 µl of 2 X Laemmli buffer. Samples were separated and analyzed by 7.5% SDS-PAGE, and then  
10 transferred to nitrocellulose membranes. The membranes were blocked with 5% bovine serum albumin (BSA) (Boehringer Mannheim Co., Indianapolis, IN) and probed with primary antibody for 1 h at RT. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) reagents (Amersham Corp., Arlington Heights, IL).

15

**SCF induces the tyrosine phosphorylation of *RAFTK* in CMK megakaryocytic cells.**

- To investigate whether *RAFTK* participated in growth-related signal transduction pathways in megakaryocytes, model CMK megakaryocytic cells were  
20 studied with a focus on the c-kit/SCF pathway which is known to be important in the proliferation of this lineage (Briddell, R. A. et al. (1991) *Blood* 78, 2854-2859; Avraham, H. et al. (1992) *Blood* 79, 365-371; Avraham, H. et al. (1992) *Blood* 80, 1679-1684). The CMK cells were starved in RPMI-1640 medium followed by stimulation with SCF, and then harvested at different time intervals as indicated. Cells  
25 were lysed and immunoprecipitated with polyclonal *RAFTK* R-4250 antibodies. The precipitates were then immunoblotted with monoclonal anti-phosphotyrosine (PY-20) antibodies or FAK monoclonal antibodies or with *RAFTK*-specific antibodies.

- Tyrosine phosphorylation of *RAFTK* peaked at a concentration of SCF of 500 ng/ml at 5 min. No phosphorylation of FAK was observed under these conditions.  
30 Maximum stimulation of *RAFTK* in CMK cells was observed within 1 min and peaked at 5 min.

- SCF treatment is known to increase cytoplasmic calcium levels (Columbo, M. et al. (1994) *Biochemical Pharmacology* 47, 2137-2145) as well as to stimulate phosphorylation of components of c-kit associated signalling pathways (Lev, S. et al.  
35 (1992) *Journal of Biological Chemistry* 267, 15970-15977; Rottapel, R. et al. (1991) *Molecular & Cellular Biology* 11, 3043-3051). Since SCF effects are modulated by PKC (Blume-Jensen, P. et al. (1995) *Journal of Biological Chemistry* 270, 14192-

14200; Namciu, S. et al. (1994) *Oncogene* 9, 1407-1416; Grabarek, J. et al. (1992) *Journal of Biological Chemistry* 267, 10011-10017), the role of PKC in *RAFTK* stimulation was investigated. SCF treatment of CMK cells induced rapid phosphorylation of *RAFTK* within 1 min and was completely blocked by the PKC inhibitors calphostin C or staurosporine.

**PMA induces tyrosine phosphorylation of *RAFTK* in CMK megakaryotic cells.**

To determine if *RAFTK* participates in pathways of megakaryocyte differentiation, the effect of PMA, which induces differentiation and maturation of cells of this lineage (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-79; Namciu, S. et al. (1994) *Oncogene* 9, 1407-1416), was studied and was found to be able to stimulate *RAFTK* phosphorylation. Under these conditions, phosphorylation of *RAFTK* peaked at 5 min at 50 nM PMA.

Because PMA activates PKC, the relationship between PMA induced tyrosine phosphorylation of *RAFTK* in CMK cells and PKC activation was further characterized. Addition of the PKC inhibitors Calphostin C (1  $\mu$ M) or staurosporine (75 nM) inhibited *RAFTK* phosphorylation following PMA treatment of CMK cells. In additional correlative experiments, PMA-sensitive isozymes were first down-regulated in CMK cells by prolonged treatment with PMA (15 min at 37°C), and then these cells were treated with SCF or PMA. This prolonged treatment with PMA completely abolished the subsequent effects of PMA or SCF treatment, suggesting that tyrosine phosphorylation of *RAFTK* by SCF or PMA is a PKC dependent mechanism. PMA stimulation did not result in an increase in the tyrosine phosphorylation of FAK, suggesting these related molecules may have distinct roles in different signalling activation pathways in megakaryocytes.

**Shc association with Grb2 in SCF stimulated CMK cells.**

The activated c-kit receptor can recruit Grb2 by tyrosine phosphorylation of Shc (Liu, L. et al. (1994) *Mol. and Cell Biol.* 14, 6926-6935; Tauchi, T. et al. (1994) *J. Exp. Med.* 179, 167-175). It was then investigated whether SCF could induce tyrosine phosphorylation of Shc and whether it associates with Grb2 in CMK cells. CMK cells stimulated by SCF (500 ng/ml) or PMA (50 nM) were lysed and immunoprecipitated with monoclonal anti-human Shc antibody and the precipitates were then immunoblotted with monoclonal anti-phosphotyrosine antibody PY-20. SCF induced Shc phosphorylation and association with Grb2 while PMA has no

effect on Shc-Grb2 association, suggesting differences in proliferative versus maturational pathways with regard to these molecules in megakaryocytes.

**PKC $\alpha$  and PKC $\delta$  isoforms are involved in *RAFTK* phosphorylation.**

5 To identify the PKC-isozymes involved in *RAFTK* phosphorylation, an analysis of PKC isozymes present in CMK cells was performed. PKC $\delta$  and PKC $\alpha$  isoforms were expressed in megakaryocytes as observed by immunoprecipitation using specific antibodies for PKC $\alpha$  and PKC $\delta$ , while no expression of PKC $\beta$  or PKC $\gamma$  was observed. To characterize which PKC-isozymes that may be involved in *RAFTK* 10 phosphorylation upon SCF or PMA stimulation of CMK cells, CMK cell lysates unstimulated or after PMA or SCF treatment were immunoprecipitated with either PKC $\delta$  or PKC $\alpha$  specific antibodies. The immunoprecipitates were resolved on 7.5% SDS-PAGE, immunoblotted with PY-20 antibodies or anti-phosphoserine/threonine or PKC $\delta$  or PKC $\alpha$  or *RAFTK* specific antibodies. Constitutive association of PKC $\delta$  15 with *RAFTK* was demonstrated by co-immunoprecipitation studies as shown in. An increase in phosphorylation of PKC $\delta$  upon PMA treatment was observed. A similar observation of augmented PKC $\delta$  phosphorylation was made in CMK cells stimulated with SCF.

PKC $\alpha$  was found only to be associated with *RAFTK* after stimulation with 20 SCF or PMA, but changes in the degree of phosphorylation of PKC $\alpha$  upon such activation by SCF or PMA using phosphoserine/threonine antibodies were not detected. These results show that PKC $\alpha$  and PKC $\delta$  isozymes are involved in *RAFTK* activation, that PKC $\delta$  is constitutively associated with *RAFTK* while association of *RAFTK* with PKC $\alpha$  is inducible, and may be increased in activation upon certain 25 stimuli but may not alter the phosphorylation status of this species.

**Association of *RAFTK* with Paxillin and PKC $\delta$ .**

To address the role of *RAFTK* in the formation of focal adhesions, *RAFTK* association with a known focal adhesion protein paxillin was investigated. Cell 30 lysates of CMK cells treated with PMA were immunoprecipitated with either *RAFTK* specific antibodies or anti-paxillin antibodies. The immunoprecipitates were resolved and immunoblotted with PY-20 antibodies. An increase in phosphorylation of paxillin was observed upon PMA stimulation which peaked by 5 min. Constitutive association of paxillin and *RAFTK* was observed in untreated CMK cells.

35 Treatment with the PKC inhibitor calphostin C abolished *RAFTK* activation and decreased its association with paxillin. Paxillin phosphorylation at 10 min was not altered by calphostin C treatment, indicating that paxillin phosphorylation is not

dependent on PKC. Similar observations of paxillin phosphorylation and its association with *RAFTK* were observed in CMK cells stimulated with SCF.

In addition, CMK cells stimulated with SCF were immunoprecipitated with *RAFTK* antibodies, and the immunoprecipitates were then resolved and  
5 immunoblotted with PY-20 antibodies, *RAFTK* or paxillin antibodies. *RAFTK* constitutively associated with paxillin and was not altered in its degree of phosphorylation upon activation with SCF. Constitutive association of PKC $\delta$  with *RAFTK* was also observed. These results demonstrate a constitutive association of PKC $\delta$ , *RAFTK* and paxillin in untreated cells. Upon activation with PMA, PKC $\delta$  was  
10 phosphorylated and peaked by 10 min. The constitutive association of PKC $\delta$  with *RAFTK* was not affected by PMA stimulation.

***RAFTK* activation is inhibited by BAPTA, calphostin C and cytochalasin-D.**

Calcium ionophore (A23187) treatment of cells elevates intracellular calcium  
15 levels and initiates a cascade of signalling events including PKC activation. CMK cells treated with calcium ionophore A23187 showed tyrosine phosphorylation of *RAFTK*, which was inhibited by the intracellular calcium chelator, BAPTA. In the presence of calphostin C, a specific PKC inhibitor, induction of *RAFTK* phosphorylation by the calcium ionophore A23187 was inhibited, indicating that  
20 calcium regulation of this PKC isoform was involved in *RAFTK* phosphorylation. Upon calcium ionophore A23187 treatment, *RAFTK* activation was completely inhibited in the presence of cytochalasin-D, indicating that *RAFTK* is associated with the cytoskeleton and this association is essential for its activation following changes in intracellular calcium.

25 Similarly, SCF treatment of CMK cells induced rapid, transient tyrosine phosphorylation of *RAFTK* which was inhibited in the presence of BAPTA, suggesting that SCF may induce *RAFTK* phosphorylation through elevating intracellular calcium levels. However, in the presence of cytochalasin-D, SCF induction of *RAFTK* phosphorylation was inhibited completely, indicating again that  
30 the integrity of the cytoskeleton is required for *RAFTK* phosphorylation. Since SCF stimulation of *RAFTK* in the presence of calphostin C was also completely inhibited, this mechanism of *RAFTK* activation appears to be mediated through PKC.

PMA phosphorylation of *RAFTK* also was blocked by calphostin C or by BAPTA, further indicating that calcium regulation of PKC isoforms is involved in  
35 *RAFTK* stimulation. Cytochalasin-D treatment inhibited PMA stimulation of *RAFTK*, suggesting that *RAFTK* association with the megakaryocytic cytoskeleton is critical for its phosphorylation in cells of this lineage.



In this study, *RAFTK* activation and its regulation in megakaryocytic cells was characterized. The results obtained in these studies demonstrated that *RAFTK*, unlike FAK, is tyrosine phosphorylated upon SCF and PMA treatments. The finding that FAK is not phosphorylated under these conditions is consistent with prior studies of FAK activation in Mo7E megakaryocytic cells (Gotoh, A. et al. (1995) *Experimental Hematology* 23, 1153-1159) and suggests important differences in the roles of FAK and *RAFTK* in cells of this lineage.

The effects of SCF, PMA and  $\text{Ca}^{2+}$  on activation of *RAFTK* were mediated through PKC. Moreover, direct association of *RAFTK* with paxillin was observed and activation of *RAFTK* resulted in a sequential activation and phosphorylation of this cytoskeletal protein.

PKC plays an important role in cellular responses to various hormones, growth factors, neurotransmitters and cytokines, and transduces signals promoting lipid hydrolysis (See, e.g., Dekker, L. V. and Parker, P. J. (1994) *Trends in Biochemical Sciences* 19, 73-77; Nishizuka, Y. (1992) *Science* 258, 607-614; Nishizuka, Y. (1986) *Science* 233, 305-312). PKC regulates the action of a variety of ion channels, G-protein coupled receptors, tyrosine kinase receptors or non-receptor tyrosine kinases (See, e.g., Ohtani, K. et al. (1995) *Journal of Neurochemistry* 65, 605-614; Rozengurt, E. (1995) *Cancer Surveys* 24, 81-96; Sadoshima, J., et al. (1995) *Circulation Research* 76, 1-15). Since the proliferative effects of SCF appeared to be modulated in part by PKC (See, e.g., Blume-Jensen, P. et al. (1995) *Journal of Biological Chemistry* 270, 14192-14200; Sato, T. et al. (1989) *Br. J. Hematol.* 72, 184-190; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231), a role for PKC in *RAFTK* stimulation was investigated. Indeed, SCF induced rapid phosphorylation of *RAFTK* and was completely blocked by the PKC inhibitors calphostin C or Staurosporine. Furthermore, PMA, which induces differentiation of CMK megakaryocytic cells (Sakaguchi, M. et al. (1991) *Blood* 77, 481-485; Cowley, S. A. et al. (1992) *Int. J. Cell Cloning* 10, 223-231; Avraham, H. et al. (1992) *Int. J. Cell Cloning* 10, 70-79; Namciu, S. et al. (1994) *Oncogen* 9, 1407-1416), also stimulated *RAFTK* tyrosine phosphorylation; this *RAFTK* phosphorylation was abolished by the PKC inhibitors or prolonged treatment with PMA. These results also demonstrate that *RAFTK* activation was mediated by PKC.

To further characterize this pathway, the role of PKC isoforms in *RAFTK* activation was studied. The PKC- $\alpha$  and PKC- $\delta$  isoforms are known to be expressed in megakaryocytes while PKC- $\gamma$  and PKC- $\beta$  have not been found (Grabarek, J. et al. (1992) *Journal of Biological Chemistry* 267, 10011-10017). Constitutive association *in vivo* of PKC- $\delta$  with *RAFTK* was observed while association of PKC- $\alpha$  and *RAFTK*

was inducible. No increase in the level of PKC- $\alpha$  or PKC- $\delta$  with *RAFTK* after SCF or PMA stimulation was observed by co-immunoprecipitation techniques. Although PKC isoenzymes do not possess intrinsic tyrosine kinase activity, activation of PKC by phorbol esters such as PMA has been demonstrated to indirectly induce tyrosine phosphorylation in different cells types (See, e.g., Li, W. et al. (1994) *Journal of Biological Chemistry* 269, 2349-2352; Einspahr, K. J. et al. (1990) *Journal of Immunology* 145, 1490-1497; Nel, A. E et al. (1990) *Journal of Immunology* 145, 971-979). In NIH-3T3 or 32D transfectants overexpressing various PKC isoenzymes, pronounced phorbol diester-dependent tyrosine phosphorylation of PKC- $\delta$  was observed, while no detectable tyrosine specific phosphorylation was found after treatment with the other PKC isoenzymes transfectants (Li, W. et al. (1994) *Journal of Biological Chemistry* 269, 2349-2352).

SCF induction of tyrosine phosphorylation of *RAFTK* could be mediated by elevated intracellular calcium levels and activation of PKC through PLC- $\gamma$  (Yeo, E. J. et al. (1994) *Journal of Biological Chemistry* 269, 27823-27826; Ma, Y. H. et al. (1994) *Journal of Biological Chemistry* 269, 30734-30739; Zirrgiebel, U. et al. (1995) *Journal of Neurochemistry* 65, 2241-2250). *RAFTK* phosphorylation was inhibited following treatment by the PKC inhibitors Calphostin C or Staurosporine as well as BAPTA (an intracellular  $\text{Ca}^{2+}$  chelator). Thus PKC is directly involved in *RAFTK* phosphorylation. BAPTA blocked PMA or SCF induced *RAFTK* phosphorylation, indicating that calcium was essential for PKC mediated *RAFTK* activation. Moreover, the calcium-ionophore (A23187) also stimulated *RAFTK* phosphorylation and was inhibited by PKC inhibitors Calphostin C and Staurosporine or by BAPTA, indicating again a role for PKC as a mediator of several signalling pathways including  $\text{Ca}^{2+}$  in *RAFTK* phosphorylation.

The cytoskeleton is essential for many cellular functions including regulation of cell shape, flexibility, and adhesive properties (Hynes, R. O. (1992) *Cell* 69, 11-25; Juliano, R. L. and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585). Part of the cytoskeleton and plasma membrane form a region known as the focal adhesion (Lo, S. H. et al. (1994) *Bioessays* 16, 817-823). Focal adhesions are structures that form adherent contacts with the extracellular matrix. Proteins contained in the focal adhesion include talin, (-actinin, vinculin, paxillin, and other proteins (See, e.g., Tachibana, K. et al. (1995) *Journal of Experimental Medicine* 182, 1089-1099; Petch, L. A. et al. (1995) *Journal of Cell Science* 108, 1371-1379; Lewis, J. M. and Schwartz, M. A. (1995) *Molecular Biology of the Cell* 6, 151-160). The signal transduction pathways initiated by integrins involves cytoskeletal dependent activation of tyrosine kinases and phosphorylation of a number of substrates including

FAK protein (See, e.g., Juliano, R. L. and Haskill, S. (1993) *J. Cell Biol.* 120, 577-585; Hamawy, M. M. et al. (1994) American Society of Microbiology, Washington DC, p. 235; Schaller, M. D. et al. (1992) *Proc. Natl. Acad. Sci. USA* 89, 5192-5196).

*RAFTK* is tyrosine phosphorylated upon fibronectin stimulation and co-localized with vinculin at "focal adhesion like structures" in CMK cells. *RAFTK* activation upon calcium ionophore (A23187) treatment or SCF or PMA stimulation of CMK cells is completely abolished in the presence of cytochalasin-D, which disrupts the cytoskeleton. These observations indicate that *RAFTK* is associated with the cytoskeleton and the integrity of the cytoskeleton is required for *RAFTK* phosphorylation, similar to that of FAK phosphorylation by integrins and external stimuli (Clark, E. A. and Brugge, J. S. (1995) *Science* 268, 233-239).

Constitutive *in vivo* association of *RAFTK* and paxillin was also observed. This effect of *RAFTK* on paxillin activation can promote paxillin binding to other SH2-domain containing proteins that might be involved in multiple signal transduction pathways.

Constitutive association between *RAFTK*, PKC- $\delta$  and paxillin *in vivo* was observed. This shows a role for *RAFTK* in linking and crosstalk between various signaling proteins localized in the cytosol and focal adhesion contacts. The RhoA-dependent assembly of focal adhesions in Swiss 3T3 cells was associated with increased tyrosine phosphorylation and the recruitment of both pp125FAK and PKC- $\delta$  to focal adhesions (Barry, S. T. and Critchley, D. R. (1994) *Journal of Cell Science* 107, 2033-2045). Association of PKC- $\delta$  with *RAFTK* and paxillin shows that phosphorylation of these components can be an important event in integrin mediated events in megakaryocytes.

The tyrosine kinase PYK2, which is identical to *RAFTK*, has been shown to be involved in calcium signalling and MAP kinases function in PC-12 neuronal cells (Lev, S. et al. (1995) *Nature* 376, 737). Stimulation of megakaryocytes with thrombin leads to tyrosine phosphorylation of *RAFTK* (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 1-10). The evidence that *RAFTK* is involved in the c-kit growth factor signalling pathway in megakaryocytic cells as described in this study further supports the broad function of this kinase in a variety of signalling pathways.

**EXAMPLE 3: Characterization of *RAFTK*, a Novel Focal Adhesion Kinase, and Its Integrin-Dependent Phosphorylation and Activation in Megakaryocytes**

Biochemical characterization and functional analysis of the *RAFTK* protein was performed. Coexpression of *RAFTK* and FAK proteins in megakaryocytic cells and blood platelets was observed. Using a specific antibody to *RAFTK* and the monoclonal antibody 2A7 to FAK, FAK and *RAFTK* could be distinguished antigenically.

*RAFTK* had intrinsic tyrosine kinase and auto-kinase activities. It was phosphorylated on tyrosine in growing cultures of COS cells transfected with the pCDNAIII/flag-*RAFTK* expression vector containing the *RAFTK* cDNA ligated with the eight amino acid flag peptide sequence. Similar to FAK, dephosphorylation of *RAFTK* was observed when adherent transfected COS cells were detached. Phosphorylation was regained upon replating of these cells on the fibronectin-coated dishes. Analysis of tyrosine phosphorylated *RAFTK* from adherent transfected COS cells revealed that the Src homology 2 (SH2) domains of the Src and Fyn protein kinases as well as the Grb2 adaptor protein were able to specifically associate with *RAFTK*. Tyrosine phosphorylation of endogenous *RAFTK* was observed upon fibronectin induced activation of human megakaryocytic cells. Furthermore, colocalization of *RAFTK* protein with vinculin, a focal adhesion protein, was observed in "focal adhesion-like structures" in adherent CMK cells and in transfected pCDNAIII/flag-*RAFTK* COS cells upon fibronectin activation, by confocal microscopy.

These data show that *RAFTK* is a novel member of the FAK family, that it localizes to "focal adhesion-like structures" in CMK megakaryocytic cells, participates in integrin-mediated signaling pathways in megakaryocytes and is able to associate with the tyrosine kinases Src and Fyn as well as the adaptor protein Grb2 via SH2-phosphotyrosine interactions.

The following materials and methods were used to biochemically characterize and functionally analyze the *RAFTK* protein:

**Chemical and biological reagents**

Human fibronectin, Poly-L-Lysine (MW 70,000-150,000 Dalton) and geneticin (G418) were purchased from Sigma Chemical Co. (St. Louis, MO).

Monoclonal antibody to phosphotyrosine (PY20) was purchased from Zymed Laboratories, Inc. (South San Francisco, CA). Monoclonal antibody 2A7 and polyclonal antibody BC3 to pp125<sup>FAK</sup> were gifts from Dr. T. Parsons (University of Virginia, Charlottesville, VA). The 2A7 antibody was also purchased from Upstate Biotechnology Inc. (Lake Placid, NY). Polyclonal antibody 331 to FAK was a gift from Dr. S.K. Hanks (Vanderbilt University, Nashville, TN). Monoclonal antibody M5 to the flag peptide was purchased from Eastman Kodak Co. (New Haven, CT). rhGM-CSF was purchased from R & D Systems (Minneapolis, MN). Monoclonal anti-human antibody to vinculin was purchased from Sigma (St. Louis, MO).

10

### Cells and cell growth

COS cells were obtained from the American Type Tissue Culture (Rockville, MD). The cells were maintained in Dulbecco's modified Eagle's medium (DMEM) (Mediatech Co., Washington, D.C.) supplemented with 10% fetal calf serum (FCS) (Sigma), 2 mM glutamine, 100 U/ml of penicillin and 100 mg/ml of streptomycin. Megakaryocytic cell lines CMK, DAMI, CMS, Meg-01 and CMK11-5 were maintained in RPMI-1640 supplemented with 10% FCS, 2 mM glutamine, 100 U/ml of penicillin and 100 mg/ml of streptomycin as described previously (Sakaguchi M et al. (1991) *Blood* 77, 481; Komatsu N et al. (1989) *Blood* 74, 42; Greenberg S M et al. (1988) *Blood* 72, 1968; Avraham H et al. (1994) *Blood* 83, 2126).

20

### Expression of GST (Glutathione S-Transferase) fusion proteins

Oligonucleotides flanking various *RAFTK* domains and containing appropriate restriction sites were synthesized. The polymerase chain reaction (PCR) was used with *RAFTK* cDNA as a template to amplify the appropriate fragments. The DNA fragments encoding amino acid (a.a) residues 26-286 (N-*RAFTK*), 375-680 (KA-*RAFTK*), 375-1009 (KC-*RAFTK*) and 681-1009 (C-*RAFTK*) of *RAFTK* were amplified by the PCR technique and the sequences for these encoding regions were confirmed by DNA sequencing. The PCR products were precleaved with BamHI and EcoRI and were ligated into the pGEX-2T expression vector (Pharmacia Biotech, Inc., Piscataway, NJ) which had also been cleaved with BamHI and EcoRI. Competent *Escherichia coli* (E. coli) DH5a were transformed, and recombinant bacterial clones were screened by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis of overexpressed fusion proteins and restriction enzyme analysis. GST-fusion proteins were produced by 1 mM isopropyl b-thiogalactopyranoside induction and purified by affinity chromatography on Glutathione-Sepharose beads (Pharmacia Biotech, Inc., Piscataway, NJ).

30

35

### Construction of pCDNAIII/flag and pCDNAIII/flag-*RAFTK*

The pCDNAIII/flag expression vector was constructed by inserting a short DNA fragment encoding a starting codon and an 8-a.a flag peptide into a pCDNAIII expression vector (Invitrogen Co., San Diego, CA) at HindIII and EcoRI sites. The two oligonucleotides used were: sense primer: 5'-AGC TTA TGG ACT ACA AGG ACG ACG ATG ACA GGG G-3' (SEQ ID NO:17); antisense primer: 5' AAT TCC CTT GTC ATC GTC GTC CTT ATG GTC CAT A-3' (SEQ ID NO:18). The cDNA encoding 1009 amino acids of human *RAFTK* was then subcloned in an EcoRI site located downstream of the flag sequences of the pCDNAIII/flag vector. The orientation and DNA sequences of the *RAFTK* cDNA were confirmed by DNA sequencing.

### Transfection of COS cells and analysis of *RAFTK* phosphorylation

COS cells were transfected by the calcium phosphate method using pCDNAIII/flag-*RAFTK* or pCDNAIII/flag expression vectors according to the manufacturer's protocol (Invitrogen Co., San Diego, CA). The transfected cells were starved in serum-free DMEM for 4-6 hr, harvested by phosphate-buffered saline (PBS) containing 2 mM EDTA and washed with PBS twice. The cells ( $1.5 \times 10^6$  per 60 mm dish) were then plated onto fibronectin (5.0  $\mu\text{g/ml}$ ) or Poly-L-Lysine (5.2  $\mu\text{g/ml}$ ) coated dishes at 37°C for various times (20 or 40 min). Adherent cells were lysed in 1 ml of RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 10  $\mu\text{g/ml}$  of aprotinin, leupeptin and pepstatin, and 1 mM  $\text{Na}_3\text{VO}_4$ ). Analysis of *RAFTK* phosphorylation was performed as described further.

### Immunoprecipitation

Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by the protein assay (Bio-Rad Laboratories, Hercules, CA) and were standardized to equal concentrations of protein prior to immunoprecipitation. Identical amounts of protein from each sample were precleared by incubation with protein G-Sepharose CL-4B (Sigma Co., St. Louis, MO) for 1 hr at 4°C. After the removal of protein G-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as described below for each experiment for 4 hr or overnight at 4°C. Immunoprecipitation of the antigen-antibody complex was accomplished by incubation for 2 hr at 4°C with 25  $\mu\text{l}$  of protein G-Sepharose. Non-specific bound

proteins were removed by washing the Sepharose beads three times with HNTG buffer (20 mM Hepes, pH 7.4, 150 mM NaCl, 10% glycerol and 0.1% Triton X-100). Bound proteins were solubilized in 20  $\mu$ l of 2 x Laemmli buffer and further analyzed by immunoblotting.

5

#### **Endogenous *RAFTK* phosphorylation upon integrin activation**

20 x 10<sup>6</sup> CMK cells were starved in serum-free RPMI-1640 culture medium overnight. The cells were divided into 4 parts: one portion was replated onto fibronectin-coated (8  $\mu$ g/cm<sup>2</sup>) dishes; one portion onto collagen-coated (8  $\mu$ g/cm<sup>2</sup>) dishes; one portion onto Poly-L-Lysine (5  $\mu$ g/cm<sup>2</sup>) dishes; and one portion was kept in suspension. After 1 hr replating, the medium was aspirated and adherent cells were gently and quickly washed with ice-cold PBS. The cells were lysed in 1 ml RIPA buffer and cleared by centrifugation for 10 min at 10,000 rpm. 800  $\mu$ g of TCL was incubated overnight at 4°C with 10  $\mu$ l of R-4250, followed by immunoprecipitation with protein-A-Sepharose beads for 2 hr. The complexes were washed with HNTG buffer 3 times and then analyzed by Western blot analysis.

15

#### **Immunoblot**

A defined amount of protein lysate was combined with the same volume of Laemmli loading buffer and boiled for 2 min. In the case of immunoprecipitates, 20  $\mu$ l of 2 x Laemmli loading buffer was added. Samples were separated and analyzed by 8% SDS-PAGE gel and then transferred to nitrocellulose membranes. The membranes were blocked with 5% bovine serum albumin (BSA) in PBS containing 0.1% Tween-20 (Boehringer Mannheim Co., Indianapolis, IN) and probed with primary antibody for 1 hr at room temperature according to the enhanced chemiluminescent (ECL) protocol (Amersham Corp., Arlington Heights, IL). Immunoreactive bands were visualized using HRP-conjugated secondary antibodies and ECL reagents (Amersham Corp., Arlington Heights, IL).

25

#### ***In vitro* kinase assay**

Immunoprecipitated complexes were washed twice with RIPA buffer and once in kinase buffer (20 mM Hepes; pH 7.4, 50 mM NaCl, 5 mM MgCl<sub>2</sub>, 5 mM MnCl<sub>2</sub>, and 100  $\mu$ M Na<sub>3</sub>VO<sub>4</sub>). The washed immune complexes or a defined amount of purified GST-fusion proteins were suspended in 20  $\mu$ l of kinase buffer and then [<sup>32</sup>P]ATP was added up to 250  $\mu$ Ci/ml in the presence of 25  $\mu$ g of poly (Glu:Tyr) (4:1, 20 to 50 Kd) (Sigma Chemical Co., St. Louis, MO), at RT for 15 min. The reaction

35

was stopped by the addition of 10 mM Hepes (pH 7.4)/10 mM EDTA. The labeled substrates were analyzed by SDS-PAGE and autoradiography.

### ***In vitro* SH2 association assays**

- 5           *In vitro* association experiments were carried out with GST-fusion proteins containing the SH2 domains of Fyn and Grb2 (a generous gift of Dr. L. Cantley, Harvard Medical School, Boston, MA) and Src (a generous gift of Dr. T. Pawson, Mount Sinai Research Institute, Canada). TCL (1 mg) was incubated for 120 min at 4°C with 10 µg of GST-fusion proteins coupled to Glutathione-Sepharose beads in the  
10           absence or presence of the indicated amount of synthetic peptide. The beads were washed three times with Tris-buffer saline, and proteins were separated by 8% SDS-PAGE.

### **Antibodies**

- 15           Anti-*RAFTK* antiserum was obtained from New Zealand white rabbits immunized with a bacterially expressed GST-fusion protein containing the C-terminal (681-1009 amino acid residues) of human *RAFTK* cDNA. The sera were titered by ELISA against the GST-*RAFTK* C-terminus fusion protein. The serum R-4250 exhibited the highest titer (1:256,000) and was found to react specifically with the  
20           GST-fusion proteins containing the C-terminus of *RAFTK*. This serum was used in subsequent experiments.

### **Confocal microscopy**

- (1) *Cell staining*: Cultured CMK megakaryocytes or transfected COS cells  
25           were plated overnight on glass coverslips coated with fibronectin (5-8 µg/cm<sup>2</sup>). Adherent cells were fixed with 2% paraformaldehyde in PBS, pH 7.4 for at least 15 min, permeabilized in PBS containing 0.2% Triton X-100 for 2 min, washed in PBS, and placed in a blocking buffer containing PBS, 3% BSA and 1% normal goat serum for 1 hr. Adherent cells were incubated with anti-vinculin (IgG1 mouse anti-human,  
30           1:200 dilution) and *RAFTK* (rabbit anti-*RAFTK*, 1:100 dilution) antibodies for 1 hr. Cells were washed in PBS, incubated with FITC conjugated (anti-mouse IgG), and Texas red conjugated (anti-rabbit IgG) antibodies for 1 hr.

- (2) *Microscopy*: Stained cells were briefly washed in PBS, then sealed in microwell chambers containing Vectashield antifade and examined using a Sarastro  
35           2000 confocal laser scanning microscope (CLSM) (Molecular Dynamics, Sunnyvale, CA) fitted with a 25 mW argon-ion laser. The microscope was configured for dual channel fluorescent imaging with: 488/514 nm excitation, 535 nm primary



beamsplitter, 10% laser transmission, 18 mW laser power. A 595 nm secondary beamsplitter passed fluorescent light emitted from *RAFTK*-labeled (Texas red)-cells to a photomultiplier tube fitted with a 600 long pass filter. Short wavelength light (<595 nm) emitted from the vinculin-labeled focal adhesion plaques (FITC) was directed to  
5 a second photomultiplier tube fitted with a  $540 \pm 15$  nm band pass filter and simultaneously recorded. Image pairs were subjected to a 2-D median filter to reduce background noise, then examined as color composite images with *RAFTK* appearing red, and vinculin green. Focal adhesion plaques expressing both *RAFTK* and vinculin appeared yellow-orange.

10 (3) *Image analysis*: In some cases, the distribution of *RAFTK* and vinculin were determined using a quantitative image analysis procedure in which 2-D pixel intensity histograms from both the *RAFTK* and vinculin images were compared using ImageSpace (Molecular Dynamics, Sunnyvale, CA) software. Image analysis was performed on pairs of images to determine the area occupied by *RAFTK* and vinculin  
15 within the CMK cells. Pixel dimensions of all micrographs were  $0.17 \mu\text{m}$ , and pixel intensities ranged from 0-255 intensity units.

The pixel intensity range corresponding to the cell cytoplasm was determined separately for each image. Cell background was found to be within a range of 4-73 pixel intensity units. The noncellular background composed of surrounding media  
20 and slide surface was determined to be 0-3 pixel intensity units. Cells containing fluorescent stained *RAFTK* were observed to have pixel intensities from 74-255, whereas pixel intensities for vinculin ranged between 86-254 units. Pixel intensities corresponding to either *RAFTK* or vinculin were applied to a 2-D histogram of pixel intensities with *RAFTK* (X-axis), plotted against vinculin (Y-axis).

25 Pixels unique to each either *RAFTK* or vinculin were identified on the 2-D histogram and converted into a binary section. This section was applied as a mask over the original image to produce area measurements on a per cell basis. This procedure was repeatedly used to measure the area occupied by both *RAFTK* and vinculin for both cell types.

30 **The *RAFTK* gene is highly conserved during species evolution and is coexpressed with FAK in several hematopoietic cells.**

*RAFTK* shares about 65% similarity of its amino acid sequence with that of FAK, suggesting that *RAFTK* and FAK may have some common features. In order to  
35 determine whether *RAFTK* and FAK may be coexpressed in the same cells, total RNA from different megakaryocytic cell lines was prepared and Northern blot analysis was performed using a human *RAFTK*-specific probe and a human FAK-specific probe.

Northern blot analysis showed that all tested megakaryocytic cell lines coexpressed both *RAFTK*- and FAK- specific transcripts of 4.5 Kb.

Coexistence of *RAFTK* and FAK proteins was detected by Western blot analysis in the CMK megakaryocytic cell line and in blood platelets using the polyclonal antibody 4250 to the C-terminal *RAFTK* and the monoclonal antibody 2A7 to FAK. Total lysates from  $5 \times 10^8$  platelets of  $10^6$  CMK cells were prepared and protein concentrations were determined. An equal amount of protein was immunoprecipitated by polyclonal antiserum 4250 for *RAFTK* or normal rabbit serum (NRS) or immunoprecipitated with the monoclonal antibody 2A7 to FAK or control irrelevant monoclonal antibody with the same isotype. The immunocomplexes were resolved by SDS-PAGE and then immunoblotted with antibody 4250 (1:1000) or immunoblotted with 2A7 (1:500). Antibodies 4250 and 2A7 were shown to be specific to *RAFTK* and FAK, respectively, as described below. Taken together, these results demonstrated that *RAFTK* and FAK were coexpressed in these hematopoietic cells.

***RAFTK* is antigenically distinct from, but related to FAK.**

In order to investigate whether *RAFTK* and FAK were antigenically related, Three DNA fragments encoding N-terminal (N-*RAFTK*), kinase catalytic (KA-*RAFTK*) and C-terminal (C-*RAFTK*) domains of *RAFTK* were subcloned into the pGEX-2T expression vector. In addition, the full length cDNA was ligated with a sequence encoding an eight amino acid flag peptide and subcloned into a pCDNAIII vector.

Expression of the GST-fusion proteins, N-*RAFTK*, KA-*RAFTK* and C-*RAFTK* as well as GST protein was analyzed by Western blot using antibodies to GST and FAK. Only the C-terminal GST-*RAFTK* (amino acids 681-1009) was recognized by polyclonal antibodies BC3 and 331, respectively, which were generated against the C-terminal of the FAK protein (Schaller M D et al. (1992) *Proc Natl Acad Sci USA* 89, 5192; Hanks S K et al. (1992) *Proc Natl Acad Sci USA* 89, 8487). This result showed that *RAFTK* is antigenically related to FAK. However, the monoclonal antibody 2A7 against FAK did not crossreact with any of the three fusion proteins, suggesting that *RAFTK* may be antigenically different from FAK.

To further confirm these findings, a polyclonal antiserum 4250 by rabbit immunization with the C-terminal *RAFTK* GST-fusion protein was generated. This antiserum, like the monoclonal antibody M5 to the flag peptide and polyclonal Ab BC3, specifically recognized the M5-immunoprecipitated flag-*RAFTK* protein expressed in transfected COS cells. To verify 2A7 antibody specificity, antibody 2A7

but not 4250 or M5, was shown to recognize a 2A7-immunoprecipitated 125 Kd protein in either vector alone or flag-*RAFTK* transfected COS cells; this protein appears to represent the endogenous FAK. Similarly, *RAFTK* immunoprecipitated by antiserum 4250 from CMK megakaryocytic cell lysates was not recognized by the monoclonal antibody 2A7 to FAK and FAK immunoprecipitated by the monoclonal antibody 2A7 from CMK cell lysates was not recognized by *RAFTK*-specific antibodies. Taken together, these data show that FAK and *RAFTK* are distinguishable antigenically while being related members of the FAK family.

**10    *RAFTK* has intrinsic protein tyrosine kinase and auto-kinase activities.**

Enzymatic activity of a protein tyrosine kinase is essential for its role in signal transduction. To assess intrinsic tyrosine kinase activity, kinase activity of the purified N-*RAFTK*, KA-*RAFTK*, KC-*RAFTK*, and C-*RAFTK* GST-fusion proteins in *in vitro* assays in which poly (Glu:Tyr) (4:1) was used as an exogenous substrate was tested. The results showed that the KC-*RAFTK* fusion protein possessed kinase activity. Similar data were obtained by testing total recombinant bacterial cell lysates, as only the KC-*RAFTK* construct had kinase activity. These results demonstrate that *RAFTK* has intrinsic PTK activity.

Autophosphorylation of FAK is an initial step in cell response to stimuli and provides a dock for the association of Src-family kinases (Weng Z et al. (1993) *J Biol Chem* 268, 14956; Schaller M D et al. (1994) *Mol Cell Biol* 14, 1680). In order to test *RAFTK* auto-kinase activity, the incorporation of <sup>32</sup>P into the purified GST-fusion proteins containing either the kinase catalytic domain, C-terminal domain or kinase plus C-terminal domain was tested in the absence of exogenous substrate. In agreement with the results from the *in vitro* kinase activity assay, only the kinase catalytic domain plus C-terminal domain of *RAFTK* appeared capable of autophosphorylating itself (MW 100 Kd). The additional low molecular weight proteins might represent the proteolytic isoforms of the KC-*RAFTK* GST-fusion protein which were found during recombinant *E. coli* growth and induction, as detected by antibodies to the GST protein. These results show that *RAFTK* possesses auto-kinase activity.

***RAFTK* is involved in integrin-mediated signaling in transiently transfected COS cells and CMK megakaryocytic cells.**

Since *RAFTK* is structurally similar to FAK which plays a central role in integrin-mediated signaling pathways, *RAFTK* phosphorylation upon integrin engagement was studied. Tyrosine phosphorylation and kinase activity of flag-

*RAFTK* in transiently transfected COS cells were analyzed. When transfected COS cells were grown on plastic culture dishes, flag-*RAFTK* protein was phosphorylated. Detachment of the transfected cells by 1 mM EDTA in PBS resulted in a significant decrease in the level of *RAFTK* phosphorylation; replating the cells onto fibronectin-coated dishes increased the phosphorylation of flag-*RAFTK* in a time dependent manner. In contrast, replating the cells onto Poly-L-Lysine coated dishes had no effect on the phosphorylation of flag-*RAFTK*. These data show that the phosphorylation of flag-*RAFTK* is modulated by integrin interaction with fibronectin.

To assess whether the kinase activity of *RAFTK* was stimulated by integrin activation, the flag-*RAFTK* from detached cells, cells attached to fibronectin or cells attached to Poly-L-Lysine were analyzed. While flag-*RAFTK* from fibronectin-coated dishes was markedly increased in its kinase activity, no increased kinase activity was found in flag-*RAFTK* from Poly-L-Lysine coated dishes. The flag-*RAFTK* in the detached cells retained a very low level of activity.

In order to elucidate whether or not endogenous *RAFTK* is responsive to integrin activation (like FAK), a detailed analysis of integrin-mediated signaling of endogenous *RAFTK* in megakaryocytes was also performed. After starvation, CMK cells were lysed in RIPA buffer as a control, or replated onto fibronectin, collagen or Poly-L-Lysine coated dishes for 1 hr. The adherent cells were quickly washed and lysed in RIPA buffer. A total of 1.2 mg TCL for each sample was immunoprecipitated with anti-*RAFTK* serum (R-4250). After washing, the immunocomplexes were divided into three parts: equivalent of 1 mg TCL for phosphorylation analysis; equivalent of 180 ug TCL for autophosphorylation assay; and equivalent of 20 ug TCL for kinase assay. The phosphorylation of *RAFTK* was significantly increased in CMK megakaryocytic cells adherent to fibronectin or collagen coated dishes, while no phosphorylation of *RAFTK* was observed in nonadherent CMK cells or CMK cells grown onto Poly-L-Lysine dishes. These data show that *RAFTK* can be activated by integrin engagement in CMK cells.

To assess whether the auto-kinase activity of *RAFTK* was stimulated by integrin activation, endogenous *RAFTK* in CMK cells untreated or stimulated with collagen, fibronectin, or Poly-L-Lysine was analyzed. There was an increase in autophosphorylation of *RAFTK* upon the collagen or fibronectin adhesion of CMK cells. In contrast, a very low level of autophosphorylation in CMK cells adherent to Poly-L-Lysine and no autophosphorylation activity in untreated CMK cells was observed.

In order to evaluate the intrinsic kinase activity of endogenous *RAFTK* in CMK cells, the kinase activity of endogenous *RAFTK* in CMK cells adherent to

fibronectin, collagen, or Poly-L-Lysine was tested. Strong phosphorylation of the exogenous substrate poly (Glu:Tyr) (4:1) by *RAFTK* was observed when CMK cells were adherent to collagen or fibronectin, while low kinase activity of *RAFTK* was detected when CMK cells were adherent to Poly-L-Lysine or in suspension. These results, together with the *in vitro* kinase experiments using the GST-*RAFTK* fusion proteins, demonstrated that *RAFTK* increased its intrinsic PTK and auto-kinase activities upon integrin activation.

**Association of phosphorylated *RAFTK* with Src-family kinases and Grb2 adaptor protein via SH2-phosphotyrosine interactions.**

The tyrosine phosphorylation of several signaling proteins provides a specific dock for the association of SH2-containing proteins. To test whether phosphorylated *RAFTK* associated with the Src-family kinases and the Grb2 protein, GST-SH2 fusion proteins from Fyn, Src, Grb2, N-terminal and C-terminal p85 subunits of PI3-K (NP85 and CP85), respectively were purified to homogeneity. These proteins were incubated with TCL from flag-*RAFTK* transfected COS cells adherent on fibronectin-coated dishes. Associated complexes were analyzed by Western blot analysis. GST-SH2 fusion proteins from Fyn, Src and Grb2 appeared to be specifically associated with phosphorylated *RAFTK* while GST protein alone, GST-NP85SH2 and GST-CP85SH2 did not bind to phosphorylated *RAFTK*. In order to test whether such association was enhanced specifically by fibronectin treatment, *RAFTK* protein from transfected COS cells in suspension was compared to that of transfected COS cells adherent on fibronectin in its association with SrcSH2, FynSH2 and Grb2SH2 GST-fusion proteins. Cells adherent on fibronectin significantly increased the association of *RAFTK* with SrcSH2, FynSH2 and Grb2SH2 GST proteins. These results demonstrate that *RAFTK* strengthens its ability to bind Src, Fyn and Grb2 molecules after cellular integrin activation.

Since Grb2SH2 GST-fusion proteins seemed to be associated less with *RAFTK* than the SrcSH2 and FynSH2 GST-fusion proteins, this specificity was further confirmed by abolishing this association with increasing concentrations of a phosphotyrosine synthetic peptide (VpYLNVMEL) corresponding to amino acids 880-887 of *RAFTK*. Taken together, these results demonstrated that tyrosine-phosphorylated *RAFTK* has the ability to specifically associate with Src and Fyn kinases and Grb2 protein in a SH2-dependent manner.

**Localization of *RAFTK* to "focal adhesion-like structures" of CMK cells and transfected COS cells.**

To analyze whether endogenous *RAFTK* in CMK cells was localized to focal  
5 adhesion structures or to cell-cell contacts, a detailed analysis using confocal  
microscopy was performed. In addition, localization of *RAFTK* in transfected COS  
cells was determined. Purified R-4250 antibodies which specifically detect *RAFTK*  
were used. Confocal image analysis of immunostained CMK cells and transfected  
10 COS cells adherent to the fibronectin substrate reveal "focal adhesion-like structures"  
adjacent to the glass coverslip surface. Confocal image analysis showed greater than  
90% of the *RAFTK* was colocalized with vinculin under these conditions. Cells  
immunostained with *RAFTK* revealed punctate areas of staining near the basal  
surfaces of cells which were adherent to fibronectin for 12 hrs. Cells adherent to  
fibronectin for 1 hr revealed similar colocalization to basal cell surfaces, however,  
15 focal adhesion plaques were not well developed at this time point.

In this study, the biochemical characterization and functional analysis of a  
novel signaling molecule, *RAFTK*, which is abundantly expressed in megakaryocytes,  
platelets and brain tissue (Avraham S et al. (1995) *J Biol Chem* 270, 27742) is  
described. The results show that *RAFTK*, like FAK, possesses intrinsic protein  
20 tyrosine kinase and auto-kinase activities; is coexpressed with FAK in  
megakaryocytic cells and platelets; and is immunologically related to, but distinct  
from FAK. The phosphorylation and kinase activity of *RAFTK* were stimulated by  
integrin engagement. Phosphorylated *RAFTK* was able to specifically bind to Src-  
family kinases and the Grb2 adaptor protein via an apparent phosphotyrosine-SH2  
25 interaction. These data demonstrate that *RAFTK* is a novel member of the FAK  
family and shares structural, immunological, enzymatic, and functional features with  
FAK.

Fibronectin stimulation increased *RAFTK* tyrosine phosphorylation when an  
epitope tagged *RAFTK* was expressed into COS cells. In addition, the  
30 phosphorylation of *RAFTK* was significantly increased in CMK cells adherent onto  
fibronectin or collagen coated dishes, while no phosphorylation of *RAFTK* was  
observed in untreated CMK cells or CMK cells grown onto Poly-L-Lysine coated  
dishes. These results clearly show that the phosphorylation of endogenous *RAFTK* is  
modulated by integrin interaction with fibronectin or collagen in CMK cells. In  
35 addition, microscopic imaging of CMK cells and transfected pCDNAIII/flag-*RAFTK*  
COS cells following double-staining with vinculin and *RAFTK* revealed  
colocalization of the *RAFTK* protein with vinculin in "focal adhesion-like structures"

in CMK and transfected COS cells treated with fibronectin. It is important to note that all published studies involving focal adhesion sites were done in adherent cells (such as 3T3 cells) where the staining of focal adhesion structures is in a conventional punctate pattern. Megakaryocytic as well as CMK cells are cells grown in suspension and lack the typical focal adhesion structures. Therefore, these confocal studies were done in transfected COS cells as well as CMK cells grown in fibronectin and the stained structures are called "focal adhesion-like structures". Furthermore, the colocalization of *RAFTK* with vinculin, a well-known focal adhesion protein, was performed in CMK and transfected COS cells. The colocalization of both proteins is about 90%. The microscopic studies support the biochemical evidence that *RAFTK* is localized to "focal adhesion-like structures" surrounded by vinculin in adherent CMK cells and transfected COS cells.

These studies on *RAFTK* localization were performed in transfected COS cells (with a flag-*RAFTK* construct) and in CMK cells. The plane of focus was set to be within the focal adhesion plaque region at 0.2 microns above the coverslip surface to exclude cell-cell contact. Furthermore, the tyrosine phosphorylation of endogenous *RAFTK* was observed upon collagen stimulation of platelets. However, in PC12 cells, *RAFTK* was not phosphorylated upon collagen treatment. These observations suggest that *RAFTK* phosphorylation upon integrin-mediated signaling is dependent on cell types and integrin forms, indicating a cell type specific signaling event.

Amino acid and DNA sequence homology studies showed that *RAFTK* is most closely related to FAK, sharing 65% similarity (Avraham S et al. (1995) *J Biol Chem* 270, 27742). Such high similarity between the proteins suggested that *RAFTK* and FAK may have similar molecular structural conformations. This prediction was supported by antigenic crossreactivity studies in which two polyclonal antibodies to FAK recognized the C-terminal GST-fusion protein of *RAFTK* as well as flag-*RAFTK*. *RAFTK* appears to be structurally distinct from FAK. A further comparison of *RAFTK* and FAK in different regions indicated that NH2 and COOH terminal domains have more divergence than the kinase domain. Such divergence may account for the failure of monoclonal antibody 2A7 to FAK and polyclonal antibody 4250 to *RAFTK* to recognize common epitopes.

Since FAK-family kinases lack SH2 and SH3 domains, the regulation of tyrosine phosphorylation plays a critical role in protein-protein interactions during signal transduction. More than half (20/35) of the tyrosine residues of *RAFTK* are highly conserved in the FAK molecule (Avraham S et al. (1995) *J Biol Chem* 270, 27742). Importantly, two of these residues in FAK were identified as being phosphorylated and sequentially bound to the SH2 domains of the Src-family kinases

and the Grb2 adaptor protein: Tyr<sup>397</sup> (Schaller M D et al. (1994) *Mol Cell Biol* 14, 1680) and Tyr<sup>925</sup> (Schlaepfer D.D. et al. (1994) *Nature* 372, 786), respectively. The sequences downstream of these two phosphotyrosines are consistent with the prediction that the Src SH2 domain preferentially binds to the phosphotyrosine sequence pYAEI, whereas Grb2 binds to pYENV (Songyang Z et al. (1993) *Cell* 72, 767). Equivalent phosphotyrosine sequences were found at Tyr<sup>402</sup> and Tyr<sup>881</sup> of *RAFTK*, however, glutamic acid next to Tyr<sup>881</sup> was substituted with leucine. These results demonstrated that the SH2 domains of Src, Fyn and Grb2 were able to specifically associate with tyrosine phosphorylated *RAFTK* from fibronectin-activated COS cells. The leucine substituted next to phosphotyrosine-881 did not change its specificity of binding for Grb2. However, unlike FAK, *RAFTK* contains more than one potential binding site for Src-family tyrosine kinases. These results show that *RAFTK*, like FAK, is a substrate for Src kinases which are required for FAK family kinase mediated signaling pathway(s).

15 Tyrosine kinase and auto-kinase activities are essential for FAK to initiate its downstream signaling pathway. Because *RAFTK* has very large NH2 and COOH domains, it is important to preclude the possibility that any other *RAFTK*-associated kinase(s) may contribute to the enzymatic activity in the *in vitro* kinase assay. Therefore, purified GST-*RAFTK* fusion proteins produced from recombinant bacteria were tested for kinase activity. The data described herein demonstrated that *RAFTK*, like FAK, possesses intrinsic kinase and auto-kinase activities. However, it is interesting to observe that such activities may require not only the kinase catalytic domain but also the COOH domain, which would differ from that observed with FAK produced in mammalian cells (Eide B L et al. (1995) *Mol Cell Biol* 15, 2819; Chan P Y et al. (1994) *J Biol Chem* 269, 20567). It is unclear whether the different observations imply different characteristics of FAK and *RAFTK* or are due to different expression systems.

30 **EXAMPLE 4: Activation of a Novel Related Focal Adhesion Tyrosine Kinase (*RAFTK*) During an Early Phase of Platelet Activation by an Integrin GpIIb-IIIa Independent Mechanism**

35 ***RAFTK* is rapidly tyrosine phosphorylated in thrombin-stimulated platelets,**  
To elucidate the role of *RAFTK* activation in platelets the effect of thrombin on *RAFTK* tyrosine phosphorylation was studied. Thrombin induced a dose and time dependent phosphorylation of *RAFTK* in platelets. Treatment of platelets with 0.05,



0.1 or 0.25 U/ml of thrombin induced a basal level of *RAFTK* phosphorylation similar to the resting or unstimulated platelets. An increase, but equal levels of *RAFTK* phosphorylation was induced with 0.5 or 1 U/ml of thrombin, while 2 U/ml of thrombin induced highest levels of *RAFTK* phosphorylation.

5 A time course of thrombin stimulation in platelets showed a rapid induction of *RAFTK* phosphorylation. The resting or unstimulated platelets showed very low basal level of phosphorylation. Activation was observed as early as 10 seconds post-thrombin stimulation, reaching a maximum at 2 minutes and tapering off by 10 minutes. These results showed that thrombin stimulation of platelets induces *RAFTK* phosphorylation in a time and concentration dependent manner.

***RAFTK* is an endogenous substrate for Calpain.**

During the course of platelet activation, agonist induced activation of calpain (Fox, J. E. B. et al. (1990) *Blood* 76, 2510-2519; Fox, J. E. et al. (1991) *J. Biol. Chem.* 266, 13289-13295; Saido, T. et al. (1993) *J. Biol. Chem.* 268, 7422-7426) and limited proteolysis of some specific substrates (Ando, Y. et al. (1987) *Biochem. Biophys. Res. Commun.* 144, 484-490; Tsujinaka, T. et al. (1982) *Thromb. Res.* 28, 149-156; Oda, A. et al. (1993) *J. Biol. Chem.* 268, 12603-12608; Frangioni, J. V. et al. (1993) *EMBO J.* 12, 4843-4856) has been reported. Calpain constitutes most of the calcium dependent protease activity in platelets (Fox, J. E. B. et al. (1990) *Blood* 76, 2510-2519; Oda, A. et al. (1993) *J. Biol. Chem.* 268, 12603-12608). Thrombin is one of the agonists that causes activation of calpain in platelets.

Despite the use of equal number of platelets ( $5 \times 10^8$  /ml) for immunoprecipitations, a dose and time dependent decrease in *RAFTK* protein levels after thrombin stimulation was consistently observed. The decrease in the protein levels could be due to protein degradation or protein processing. Since *RAFTK* undergoes processing upon thrombin stimulation, an examination was performed to determine if calpain is involved in this process by using a specific membrane permeable calpain inhibitor, calpeptin (Tsujinaka, T. et al. (1988) *Biochem. Biophys. Res. Commun.* 153, 1201-1208). When platelets were activated by thrombin, no differences were seen in the levels of *RAFTK* phosphorylation between calpeptin pretreated or thrombin alone treated platelets. There was no detectable phosphorylation in untreated or resting platelets. However, when *RAFTK* protein levels were examined on the same immunoblot, thrombin treated sample showed cleavage of *RAFTK* where as calpeptin treated platelets showed a complete blockage in *RAFTK* degradation. The level of the *RAFTK* protein in calpeptin treated platelets was equal to the level in the resting or unstimulated platelets. Furthermore, *RAFTK*

processing was observed in a time dependent manner in response to the pharmacological activator of calpain, A23187, and to the physiological activator of calpain collagen. The characteristics of inhibition of degradation of *RAFTK* were consistent with the involvement of calcium dependent neutral cysteine protease, since  
5 degradation of *RAFTK* occurred at neutral pH and was inhibited by cysteine protease inhibitor, calpeptin. These results provide evidence that agonist induced *RAFTK* processing is mediated through the activation of calpain.

10 ***RAFTK* phosphorylation is independent of aggregation and occurs early, during platelet shape change.**

The earliest platelet response induced by physiological agonists involves a change in shape from flat discs into compact spheres followed by secretion of granular contents. The later phase of platelet response is aggregation when large platelet aggregates are formed. Since *RAFTK* phosphorylation is induced early (10 seconds)  
15 after thrombin stimulation, it was investigated if aggregation is a prerequisite for *RAFTK* phosphorylation. In order to prevent aggregation, platelets were activated by thrombin in the absence of stirring. A time course of thrombin induced platelets showed that *RAFTK* is rapidly phosphorylated as early as 10 seconds in the presence or absence of stirring (aggregation). However, the signal declined faster upon stirring,  
20 where as it persisted longer in the absence of stirring (10 min). The levels of *RAFTK* protein remained equal in the absence of stirring and the levels decreased only in the presence of stirring, indicating that *RAFTK* cleavage is dependent on aggregation, while *RAFTK* phosphorylation does not require aggregation. However some aggregation can occur in the absence of stirring because of the close proximity of the  
25 platelets ( $5 \times 10^8/\text{ml}$ ). Therefore, in order to further confirm that *RAFTK* phosphorylation does not require aggregation, platelets were pretreated with an aggregation inhibitor, RDGS tetrapeptide, before activation by an agonist. Treatment with RGDS peptide allows shape change but prevents aggregation by blocking the interaction of fibrinogen with the integrin GpIIb/IIIa. Addition of thrombin to  
30 platelets pretreated with increasing concentrations of RGDS or RGES (mock) peptides showed no change in the levels of tyrosine phosphorylation of *RAFTK*. Thus, these results confirm that *RAFTK* phosphorylation does not require aggregation and it occurs during platelet shape change. Furthermore, consistent with previous findings, proteolytic processing of *RAFTK* was inhibited by RGDS treatment (absence of  
35 aggregation), but it was not inhibited when platelets were treated by RGES (presence of aggregation) or with thrombin alone. The phosphorylation of *RAFTK* correlated well with an early wave of tyrosine phosphorylation especially of proteins pp60<sup>src</sup>

and pp72<sup>syk</sup> (Clark, E. A., and Brugge, J. S. (1993) *Mol. Cell. Biol.* 13, 1863-1871; Clark, E.A. et al. (1994) *J. Biol. Chem.* 269, 288859-28864), but it precedes the wave of tyrosine phosphorylation of its family member, pp125<sup>FAK</sup> known to be dependent on platelet aggregation (Clark, E. A. et al. (1994) *J. Biol. Chem.* 269, 288859-28864).

**Activation of *RAFTK* is not dependent on the integrin, GpIIb/IIIa.**

GpIIb/IIIa is a major integrin receptor which plays an important role in adhesive events critical in clot formation by binding to fibrinogen and von Willebrand factor matrix proteins (Fox, J. E. B. et al. (1993) *J. Biol. Chem.* 268, 25973-25984). Activation of FAK was found to be mediated through the integrin GpIIb/IIIa (Lipfert, L. et al. (1992) *J. Cell Biol.* 119, 905-912). Since *RAFTK* is a member of FAK subfamily, whether *RAFTK* and FAK have similar mechanisms of regulation was investigated. Tyrosine phosphorylation of *RAFTK* was studied under conditions that specifically induce or inhibit fibrinogen binding to this receptor. The monoclonal antibody 7E3 binds to GpIIb/IIIa and blocks fibrinogen binding (Coller, B. S. et al. (1989) *Circulation* 80, 1766-1774). Incubation of platelets with 7E3 for 20 minutes prior to stirring, followed by addition of thrombin did not inhibit tyrosine phosphorylation of *RAFTK* while tyrosine phosphorylation of FAK was inhibited under the same conditions. Pretreatment with a (control) monoclonal antibody 6D1, specific for collagen receptor GpIa/IIb did not alter thrombin induced phosphorylation of *RAFTK* or FAK. Preincubation of platelets with 7E3, 6D1 or buffer alone without thrombin stimulation showed no phosphorylation of *RAFTK* or FAK. These results showed that phosphorylation of *RAFTK* was not dependent either on fibrinogen binding to GpIIb/IIIa or platelet aggregation, and therefore, the phosphorylation of *RAFTK* in platelets is not solely regulated through ligand occupancy of the integrin GpIIb/IIIa. Furthermore it is interesting to note that preincubation with 7E3 plus thrombin stimulation, but not 6D1 prevented proteolytic processing of *RAFTK*. These results indicate that *RAFTK* phosphorylation is not dependent on GpIIb/IIIa, but the proteolytic processing of *RAFTK* is dependent GpIIb/IIIa. To further confirm that phosphorylation of *RAFTK* is not dependent on the activation (crosslinking) of GpIIb/IIIa, fibrinogen binding to the stirred platelets was initiated by an anti-B3 antibody Fab fragment (anti-LIBS6), in the absence of an agonist. This antibody renders GpIIb/IIIa competent to bind fibrinogen, but it does not itself cause detectable platelet activation (Huang, M.-M. et al. (1993) *J. Cell Biol.* 122, 473-483). *RAFTK* was not phosphorylated when platelets were stirred or unstirred with anti-LIBS6 and fibrinogen, in platelets treated with fibrinogen alone or in resting platelets despite

containing the protein. *RAFTK* was however activated when platelets were treated with thrombin (positive control). Thus these studies indicate that activation of *RAFTK* does not require crosslinking of GpIIb/IIIa receptors on the platelet surface.

5 ***RAFTK* activation is regulated by actin polymerization.**

Thrombin stimulation in platelets leads to actin polymerization and causes dramatic rearrangements of the cytoskeleton thereby inducing the formation of focal-contact like areas (Furman, M. I. et al. (1993) *Thromb. Haemostasis* 70, 229-232). It was also examined whether phosphorylation of *RAFTK* was affected by agents that  
10 disrupt the actin cytoskeleton. Platelets pretreated with cytochalasin D block agonist induced actin polymerization but do not inhibit platelet aggregation. Pretreatment with cytochalasin D inhibited tyrosine phosphorylation of *RAFTK* in thrombin stimulated platelets. However the level of inhibition was not 100%, suggesting that the actin-dependent cytoskeletal interactions effected partially phosphorylation of  
15 *RAFTK*. Furthermore proteolytic processing of *RAFTK* is not inhibited in cytochalasin D treated platelets.

***RAFTK* is activated by multiple platelet agonists.**

*RAFTK* has been found to be activated by thrombin, calcium ionophore,  
20 collagen, and the combination of ADP plus epinephrine.

**EXAMPLE 5: Identification of a Novel Signal Transduction Pathway in Human Macrophages Mediated by the Related Adhesion Focal Tyrosine Kinase (*RAFTK*)**  
25

*RAFTK*, a novel non-receptor protein kinase, has been shown to be involved in focal adhesion signal transduction pathways in neuronal PC12, megakaryocytes and platelets. Because focal adhesions may modulate cytoskeleton function and thereby  
30 alter phagocytosis, cell migration, and adhesion in macrophages, the role of *RAFTK* signaling in these cells was investigated. *RAFTK* was abundantly expressed in THP1 monocytic cells as well as primary alveolar and peripheral blood derived macrophages. Phorbol diester stimulation of THP1 cells increased tyrosine phosphorylation of *RAFTK* by 2.5 minutes. Similar increases in phosphorylation were  
35 detected within 1 minute after CSF-1/MCSF stimulation. While *RAFTK* was phosphorylated with similar kinetics in peripheral blood derived macrophages, alveolar macrophages showed high constitutive phosphorylation levels which

decreased over increased time after treatment with either PMA or CSF-1/M-CSF. Immunoprecipitation analysis identified constitutive associations between *RAFTK* and the cytoskeleton protein paxillin and the signaling molecule PI-3 kinase. However, both these molecules appear disassociate from *RAFTK* at the peak time of phosphorylation after PMA or CSF-1/M-CSF stimulation. *RAFTK* was also found to preferentially associate with the amino terminus-SH3 domain of the Grb2 adaptor protein in THP1 cells. Furthermore the CSF-1/M-CSF receptor *fms* and *RAFTK* appear to associate in response to CSF-1/M-CSF treatment of THP1 cells. These data demonstrate that *RAFTK* participates in macrophage signal transduction pathways mediated by CSF-1/M-CSF.

With this background, it was investigated whether *RAFTK* was expressed in human monocyte-macrophages and whether it participated in CSF-1/M-CSF induced signaling. In parallel, the effects of treatment of monocyte-macrophages with the known chemical activator phorbol diester PMA were tested. It was observed that *RAFTK* was robustly expressed in both peripheral blood derived monocyte-macrophages as well as tissue derived alveolar macrophages. Moreover, it was activated upon treatment of mononuclear phagocytes with CSF-1/M-CSF or PMA and associated with other signaling molecules and the cytoskeletal protein paxillin. These observations provide new data on CSF-1/M-CSF signaling and molecules which may contribute to focal adhesion formation in cells of this lineage.

The following materials and methods were used to identify the novel signal transduction in human macrophages mediated by *RAFTK*:

## Cells and cell cultures

The permanent human monocyte-macrophage cell line THP-1 was obtained from the American Type Culture Collection (ATCC) and shown to be mycoplasma free prior to expansion in culture. The cells were carried in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% fetal calf serum (FCS), 2mM glutamine, sodium pyruvate, non-essential amino acids and 50µg/ml penicillin and 50µg/ml streptomycin. Primary human peripheral blood monocyte-macrophages (MMs) were obtained by phlebotomy of normal volunteers after obtaining informed consent and isolated by ficoll hypaque density centrifugation as previously described (Boyum, A. (1968) *Scand. J. Lab Invest.* 21, (Suppl. 97). Mms plated on plastic dishes for 24 hours were shaken at 200 RPM for 15 minutes and washed 3X with HBSS to remove non-adherent cells. The adherent population of cells used for subsequent studies as previously described (Kharbanda, S. et al. (1995) *Proc. Nat.*

*Acad. Sci. USA*, 92, 6132-6136). Alveolar macrophages (AMs) were obtained by bronchoalveolar lavage of normal non-smoking volunteers after informed consent was obtained and using standard procedures.

## 5 Reagents and materials

Phorbol 12-myristate 13-acetate (PMA) was obtained from Sigma (St. Louis, MO) and dissolved in dimethyl sulfoxide and stored at -20°C until use. Recombinant human CSF-1/M-CSF was kindly provided by Genetics Institute.

10 The monoclonal antibodies against phosphotyrosine (4G10), PI-3 kinase p85 regulatory subunit, Grb2, and paxillin and the polyclonal rabbit antisera antibody to the human c-fms receptor were obtained from Upstate Biotechnology. Specific polyclonal antibodies to *RAFTK* were generated by immunizing New Zealand White rabbits with a bacterially expressed fusion protein consisting of GST and the carboxy terminus (amino acids 681-1009) of human *RAFTK* cDNA subcloned into the pGEX-  
15 2T expression vector as described (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751). High titer *RAFTK* antiserum (R-4250)) was employed in subsequent experiments and was shown to be specific and not cross reactive with FAK in prior experiments (Avraham, S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751).

The Grb2 and PI-3 kinase GST fusion proteins were obtained from Santa  
20 Cruz Biotechnology. Electrophoresis reagents and nitrocellulose membranes were obtained from Bio-Rad Laboratories (Hercules, CA). All other chemicals including the protease inhibitors pepstatin, antipain, chymostatin, leupeptin, aprotinin, and alpha 1 antitrypsin were obtained from Sigma (St. Louis, MO). Because endotoxin is known to alter monocyte-macrophage function, all media and reagents were shown to  
25 be free of endotoxin contamination by Limulus endotoxin assay (Sigma Chemical) prior to using in cell cultures.

## Signal transduction studies

Cells were initially starved in DMEM with 0.5% FCS and stimulated in HBSS  
30 at density of  $5 \times 10^6$ /ml for varying time periods at 37°C with PMA (2nM-200nM/ml) or CSF-1/M-CSF (10U-10,000U/ml). After stimulation  $20 \times 10^6$  cells were microfuged for 10 seconds and lysed in 1 ml of ice cold modified RIPA buffer (50mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150mM NaCl, 1mM PMSF, 10 ug/ml of pepstatin, antipain, chymostatin, leupeptin, aprotinin, alpha 1  
35 antitrypsin, 10mM sodium fluoride and 10 mM sodium pyrophosphate). Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by protein assay (BioRad Laboratories).

### Immunoprecipitation and Western blot analysis

For immunoprecipitation studies, identical amounts of protein from each sample were clarified by incubation with protein A-Sepharose CL-4B (Pharmacia Biotech) for 1 h at 4°C. Following the removal of protein A-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as detailed below for each experiment for 4 h or overnight at 4°C. Immunoprecipitations of the antibody-antigen complexes were performed by incubation for 2 h at 4°C with 75 µl of protein A-Sepharose (10% suspension). Non-specific bound proteins were removed by washing the Sepharose beads three times with modified RIPA buffer and one time with Phosphate buffered saline (PBS). Bound proteins were solubilized in 30 µl of 2 X Laemmli buffer and further analyzed by immunoblotting. Samples were separated on 7.5% SDS-PAGE and then transferred to nitrocellulose membranes. The membranes were blocked with 5% non-fat milk protein and probed with primary antibody for 3 h at RT or 4°C overnight. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) system (Amersham Corp., Arlington Heights, IL). Blots were stripped (2% SDS, 62.5mM Tris, 100mM Beta Mercaptoethanol) for 30 minutes at 50°C and washed in TBS-T for 60 minutes before blocking and re-probing with primary antibodies.

### GST-fusion protein binding studies

GST-fusion protein Grb2-SH3 N-terminal domain, Grb2-SH3 C-terminal domain, Grb2-SH2 domain, and PI-3 Kinase-SH3 domain of the p85 regulatory subunit were purchased from Santa Cruz Biotechnology (Santa Cruz, CA). For the binding experiments, 1 mg of cell lysate was mixed with 5 g of GST-fusion protein and incubated for 1 h at 4°C on a rotatory shaker. 50 µl of glutathione Sepharose 4B beads (Pharmacia Biotech) were added to preabsorb the complex. Following incubation for 3 h at 4°C on a rotatory shaker, the beads were centrifuged and washed three times with modified RIPA buffer. The bound proteins were eluted by boiling in Laemmli sample buffer and subjected to SDS-PAGE on 7.5% gel and Western Blot analysis.

### *RAFTK* is expressed and phosphorylated in human monocyte-macrophages.

To further characterize signaling pathways in human MMs involved in their growth, differentiation and function, the permanent monocyte-macrophage cell line THP-1 as well as primary peripheral blood derived MMs or AMs were used as a model. Analysis by immunoprecipitation revealed abundant *RAFTK* protein in these

cells. There appeared to be low levels of constitutive phosphorylation of *RAFTK* in these cells under unstimulated culture conditions.

Then it was addressed whether various stimuli associated with mononuclear phagocyte activation modulated *RAFTK* phosphorylation. Preliminary experiments  
5 determined that 1000U/ml was optimal for CSF-1/M-CSF and 20nM/ml was optimal for PMA stimulation of *RAFTK* in THP1 cells and primary macrophage cultures. An increase in the tyrosine phosphorylation of *RAFTK* was specifically observed in THP1 cells following PMA and CSF-1/M-CSF treatment. The membrane was then stripped and reprobbed with anti-*RAFTK* antibody to confirm that equivalent amounts of  
10 *RAFTK* were loaded in each lane.

To determine the time course of tyrosine phosphorylation of *RAFTK*, THP1 cells were stimulated with PMA or CSF-1/M-CSF. Phosphotyrosine levels in *RAFTK* immunoprecipitates peaked at 2.5 minutes and declined by 5 minutes. However, phosphotyrosine levels again increase by 7.5 minutes and decline after 10 minutes.  
15 There were not any changes in *RAFTK* protein levels to explain these fluctuations in phosphotyrosine levels. The membrane was then stripped and reprobbed with anti-*RAFTK* antibody to confirm that equivalent amounts of *RAFTK* were loaded in each lane.

CSF-1/M-CSF treatment of THP1 cells resulted in maximum tyrosine  
20 phosphorylation of *RAFTK* within 0.5 minutes to 1 minute which declined by 2.5 minutes. Similar to the case following PMA stimulation, the phosphotyrosine levels on *RAFTK* also appeared to increase by 10 minutes. Longer stimulation times confirmed the fluctuation of *RAFTK* tyrosine phosphorylation.

CSF-1/M-CSF stimulation of MM resulted in peak tyrosine phosphorylation  
25 on *RAFTK* by 1 minute which gradually decreased over time. *RAFTK* in AM appeared to have a high constitutive degree of phosphorylation which, in response to CSF-1/M-CSF stimulation, increased slightly by 1 minute and gradually decreased over time. Anti-*RAFTK* immunoblotting of *RAFTK* immunoprecipitates showed the ~120Kd phosphoprotein corresponded to the *RAFTK* protein. Depending on the  
30 resolution of the gels, *RAFTK* was seen to migrate either as a single band or as a doublet.

*RAFTK* associates with the signaling molecules PI-3 kinase and Grb-2. Because *RAFTK*, like FAK, acts as a platform kinase site for the coalescence of signaling and adaptor molecules at sites of focal adhesions, *RAFTK* immunoblots  
35 were examined for associating co-precipitating proteins. A specific association of *RAFTK* with PI-3 kinase, an important enzyme in modulating of phosphoinositol signaling (Auger, K. R., and Cantley, L. C. (1991) *Cancer Cells* 3, 263-270) was



observed. This association was confirmed by incubating THP1 cell lysates using a PI-3 kinase GST-fusion protein, immunoprecipitated with glutathione-conjugated beads and detecting the bound proteins by anti-*RAFTK* immunoblotting. Time course studies using either CSF-1/M-CSF or PMA treatment demonstrated that the PI-3 kinase/*RAFTK* association fluctuates. While the PI-3 kinase signal appears to weaken after one minute its association with *RAFTK* strengthens after 5 minutes of stimulation. A similar pattern is detected with longer stimulation times.

Because *RAFTK* has been shown to associate with various SH2 and SH3 domain-containing proteins, the ability of *RAFTK* to form *in vitro* complexes with the adaptor molecule Grb2 was then examined. Grb2 is generally capable of associating with signaling molecules either through one of its two SH3 or through its SH2 domain. Studies were thus performed to determine which of the Grb2 regions may mediate its interaction with *RAFTK*. THP1 cell lysates incubated with GST fusion proteins corresponding to the amino terminus SH3, the SH2 and the carboxyl terminus SH3 were immunoprecipitated with glutathione-conjugated beads and the bound proteins were detected by anti-*RAFTK* immunoblotting. *RAFTK* strongly associates with the Grb2 amino terminus-SH3 domain. There was some intermediate interactions between the SH2 domain and *RAFTK* molecule in CSF-1/M-CSF stimulated THP1 cells.

***RAFTK* associates with the c-fms receptor upon mononuclear phagocyte cell activation with CSF-1/M-CSF.**

Because CSF-1/M-CSF stimulation of THP1 cells or primary macrophages appeared to have very rapid effects on *RAFTK* phosphorylation, whether *RAFTK* may directly associate with the c-fms receptor was examined. A specific association of *RAFTK* with the c-fms receptor upon CSF-1/M-CSF treatment of cells was observed. Associations were detected in reciprocal blotting experiments of THP1 cell lysates either immunoprecipitated with *RAFTK* followed by c-fms immunoblot or c-fms immunoprecipitation followed by *RAFTK* immunoblot. No association was detected between these molecules in unstimulated or PMA stimulated THP1 cells.

***RAFTK* associates with the cytoskeletal protein paxillin in mononuclear phagocyte.**

Following the observation that *RAFTK* may be coimmunoprecipitated with molecules previously characterized as components of MM signaling pathways, it was desirable to determine whether certain cytoskeletal molecules in such cells also associate with this novel kinase. Using specific antibodies to *RAFTK* or paxillin, an

important molecule involved in focal adhesions, a clear association of these two molecules in THP1 cells was found. Similar to PI-3 kinase, the strength of paxillin's constitutive association with *RAFTK* transiently fluctuated after 1 minute stimulation by either CSF-1/M-CSF or PMA stimulation. The paxillin *RAFTK* association  
5 appeared to return to constitutive levels after 5 minutes stimulation.

These studies indicate that human mononuclear phagocytes, including peripheral blood derived MMs and tissue derived AMs, express *RAFTK*, a recently identified signaling molecule that is a member of the FAK family. *RAFTK* appeared to participate in certain previously described signaling pathways following activation  
10 of these cells. Treatment with CSF-1/M-CSF and phosphorylation of the cognate c-fms receptor revealed robust phosphorylation of *RAFTK* in both the model THP-1 cell line as well as in primary macrophages. Parallel studies using the chemical activator PMA also revealed phosphorylation of *RAFTK* in macrophages in a time and concentration dependent manner.

The phosphorylation *RAFTK* was found to result in association with several well characterized components of cell signal pathways, including the enzyme PI-3 kinase and the adaptor molecule Grb2. Grb2 is an adaptor protein that has the capacity to link with a number of kinases and substrates and functions to facilitate signaling through creation of physical associations of such partners in enzymatic  
20 reactions (Pawson, T. (1995) *Nature* 373, 573-580). PI-3 kinase appears to modulate phosphoinositol metabolism in a variety of cell types, including mononuclear phagocytes, and appears to be an important component of tyrosine kinase-regulated signaling pathways that lead to cell proliferation (Gold, M. R. et al. (1994) *J. Biol. Chem.* 269, 5403-5412). CSF-1/M-CSF has been reported to induce the direct  
25 association of the p85a subunit of PI-3 kinase with the SH2 domain of Grb2 and Grb2-Sos complexes, supporting its role upstream of the Ras signaling pathway in monocytes (Saleem, A., et al. (1995) *J. Biol. Chem.* 270, 10380-10388). In addition, PI-3 kinase activation and the production of its metabolites has been suggested to be an upstream activator of calcium-independent form of PKC (Herrera-Velit, P. and  
30 Reiner, N. E. (1996) *J. Immunology* 156, 1157-1165).

These observations on *RAFTK* show that this recently identified signaling molecule plays a variety of roles in transduction of MM signaling. It provides a missing link in prior studies of CSF-1/M-CSF induced integrin expression and the subsequent formation of focal adhesion contacts, as reported by De Nichilio and  
35 Yamada (De Nichilo, M. D. and Yamada, K. M. (1996) *J. Biol. Chem.* 271, 11016-11022).

Taken together, the data presented herein contributes to the model of CSF-1/M-CSF mediated signaling in mononuclear phagocytes. c-Fms has been reported to form associations with Shc, Grb2 and Sos1 in myeloid cells suggesting c-Fms signals through the Ras pathway (Liobin, M. N. et al. (1994) *Molecular and Cellular Biology* 14, 5682-5691). The finding that *RAFTK* associated with c-Fms in CSF-1/M-CSF stimulated THP1 cells shows that c-Fms also signals through focal adhesion contacts in concert with integrin binding. Thus, macrophages, like megakaryocytes, prominently utilize *RAFTK* in cytokine mediated pathways of activation that are linked to focal contact formation. The confluence of *RAFTK* and other kinases and cytoskeletal molecules provides a platform for the interactions of signaling molecules and adaptor proteins that regulate cell morphology to finely control certain components of the immune response such as adhesion or migration.

**EXAMPLE 6:        *RAFTK*, a Novel Member of the Focal Adhesion Kinase Family, Participates in T-Cell Receptor Signal Transduction**

It has been found that *RAFTK* is constitutively expressed in human T-cells and is rapidly phosphorylated upon the activation of the T-cell receptor (TCR) for antigen. This activation results in an increase in the association of *RAFTK* with the Src cytoplasmic tyrosine kinase Fyn. *RAFTK* also associates with the SH2 domain of Grb2 and with the cytoskeletal protein paxillin. The tyrosine phosphorylation of *RAFTK* following T-cell receptor-mediated stimulation was reduced by the pre-treatment of cells with cytochalasin D, indicating the role of the cytoskeleton in this process. These observations show that *RAFTK* participates in T-cell receptor signaling and acts to link signals from the cell surface to the cytoskeleton and thereby effect the host immune response.

It was observed that *RAFTK* is phosphorylated in response to the activation of certain integrins in megakaryocytes (Li, J. et al. (1996) *Blood* 88, 417-428) and B-lymphocytes. The induced phosphorylation of *RAFTK* via calcium-mediated ion channel pathways was shown first in PC-12 pheochromocytoma cells (Lev, S. et al. (1995) *Nature* 376, 737-745) and subsequently in megakaryocytes.

Thus, *RAFTK* is expressed in human T-lymphocytes and participates in signaling events triggered by the ligation of the TCR/CD3 complex. Several of the interacting molecules that associate with *RAFTK* in human T-cells, including the cytoskeletal protein paxillin have been characterized.

The following materials and methods were used to determine the participation of *RAFTK* in T-cell receptor signal transduction:

### Cells and cell cultures

- 5           The permanent human T-cell lines Jurkat and H9 were obtained from the American Type Culture Collection (ATCC) and shown to be mycoplasma-free prior to their expansion in culture. The cells were carried in Dulbecco's Modified Eagle's Medium (DMEM) with 10% fetal calf serum (FCS), 2 mM glutamine, 50 g/mL penicillin and 50 g/ml streptomycin. Primary human peripheral blood lymphocytes
- 10 (PBLs) were obtained by phlebotomy of normal volunteers after obtaining their informed consent and isolated by ficoll hypaque density centrifugation as previously described (Boyum, A. (1968) *Scand. J. Lab Invest.* 21, (Suppl. 97)). The anti-CD3 producing hybridoma (OKT-3) was obtained from ATCC and grown in Iscove's modified Dulbecco's medium with 20% FCS. For antibody production, cells were
- 15 grown in serum-free and protein-free hybridoma medium (Sigma, St. Louis, MO) containing Nutridoma-HU 1% (Boehringer Mannheim, Indianapolis, IN).

### Reagents and materials

- 20           The lectin phytohemagglutinin (PHA) was obtained from Pharmacia Biotech (Piscataway, NJ). The nitrocellulose membrane was obtained from Bio-Rad Laboratories (Hercules, CA). The anti-phosphotyrosine monoclonal antibody (4G10, IgG2a) was obtained from Upstate Biotechnology Inc. (Lake Placid, NY). The anti-CD3 antibody X35 was obtained from Immunotech (Marseille, France), and OKT-3 was purified from OKT-3 producing hybridoma supernatants on protein A-Sepharose
- 25 columns. Antibodies to Fyn were obtained from Santa Cruz Biotechnology (Santa Cruz, CA) and Transduction Laboratories (Lexington, KY). Anti-paxillin antibody was obtained from ICN Biomedical Inc. (Costa Mesa, CA). Specific polyclonal antibodies to *RAFTK* were generated by immunizing New Zealand White rabbits with a bacterially expressed fusion protein consisting of GST and the carboxy terminus
- 30 (amino acids 681-1009) of the human *RAFTK* cDNA subcloned into the pGEX-2T expression vector as described (Li, J. et al. (1996) *Blood* 88, 417-428). The sera were titrated against the GST-*RAFTK* C-terminus fusion protein by ELISA and the serum (R-4250) which revealed the highest titer (1:256,000) was employed in the subsequent experiments. This antiserum was shown to be specific and not crossreactive with
- 35 FAK in prior experiments (Li, J. et al. (1996) *Blood* 88, 417-428). Electrophoresis

reagents were obtained from Bio-Rad Laboratories (Hercules, CA). The protease inhibitors leupeptin, aprotinin, and alpha 1 antitrypsin and all other reagents were obtained from Sigma Co. (St. Louis, MO).

## 5 Stimulation of cells

Cells were washed twice with Hanks Balanced Salt Solution, Gibco (Grand Island, NY) and resuspended at  $5 \times 10^6$  cells/ml in DMEM medium. Cells were stimulated with either PHA (10 g/ml) or CD3 antibodies X35 (10 g/ml) or OKT-3 (10 g/ml) at 37°C for various time periods. In some experiments, cells were pretreated with cytochalasin D (2  $\mu$ M) for 60 min at 37°C before stimulation. After stimulation,  $20 \times 10^6$  cells were microfuged for 10 seconds and lysed in 1 ml of modified RIPA buffer (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 10 g/ml of aprotinin, leupeptin and pepstatin, 10 mM sodium vanadate, 10 mM sodium fluoride and 10 mM sodium pyrophosphate). Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by protein assay (Bio-Rad Laboratories).

## Immunoprecipitation and Western blot analysis

For immunoprecipitation studies, identical amounts of protein from each sample were clarified by incubation with protein A-Sepharose CL-4B (Pharmacia Biotech) for 1 h at 4°C. Following the removal of protein A-Sepharose by brief centrifugation, the solution was incubated with different primary antibodies as detailed below for each experiment for 4 h or overnight at 4°C. Immunoprecipitations of the antibody-antigen complexes were performed by incubation for 2 h at 4°C with 75  $\mu$ l of protein A-Sepharose (10% suspension). Non-specific bound proteins were removed by washing the Sepharose beads three times with modified RIPA buffer and one time with Phosphate buffered saline (PBS). Bound proteins were solubilized in 40  $\mu$ l of 2 X Laemmli buffer and further analyzed by immunoblotting. Samples were separated on 7.5% SDS-PAGE and then transferred to nitrocellulose membranes. The membranes were blocked with 5% non-fat milk protein and probed with primary antibody for 3 h at room temperature or 4°C overnight. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) system (Amersham Corp., Arlington Heights, IL).

## 35 GST-fusion protein binding studies

GST-fusion protein Grb2-SH3 N-terminal domain, Grb2-SH3 C-terminal domain, Grb2-SH2 domain, and Fyn-SH2 and -SH3 domains were purchased from

Santa Cruz Biotechnology (Santa Cruz, CA). For the binding experiments, 1 mg of cell lysate was mixed with 5 g of GST-fusion protein and incubated for 1 h at 4°C on a rotatory shaker. 50 µl of glutathione Sepharose 4B beads (Pharmacia Biotech) were added to preabsorb the complex. Following incubation for 3 h at 4°C on a rotatory shaker, the beads were centrifuged and washed three times with modified RIPA buffer. The bound proteins were eluted by boiling in Laemmli sample buffer and subjected to SDS-PAGE on 7.5% gel and Western Blot analysis.

**RAFTK is expressed in human T-lymphocytes and is phosphorylated upon T-cell activation.**

To further characterize signaling pathways in human T-cells involved in the immune response, Two permanent T-cell lines, Jurkat and H9, were utilized as well as primary circulating PBLs. Analysis by immunoblot or immunoprecipitation revealed abundant *RAFTK* protein in these cells.

The stimulation of human T-cell lines with T-cell receptor ligation induces the tyrosine phosphorylation of a phosphoprotein around 115 Kd (Motto, D. G. et al. (1994) *Journal of Biological Chemistry* 269, 21608-21613; Hsi, E. D. et al. (1988) *The Journal of Biological Chemistry* 264, 10836-10842). It was investigated whether various stimuli associated with such T-cell activation modulated the phosphorylation of *RAFTK*, which has a deduced molecular weight of ~120 Kd. An increase in the tyrosine phosphorylation of *RAFTK* could be specifically observed in the T-cell lines Jurkat or H9 following T-cell receptor ligation or treatment with the lectin PHA. The membrane was then stripped and reprobed with anti-*RAFTK* antibody to confirm that equivalent amounts of *RAFTK* were loaded in each lane. Stimulation of primary circulating PBLs with anti-T-cell receptor antibody also induced an increase in the tyrosine phosphorylation of *RAFTK*.

To determine the time course of tyrosine phosphorylation of *RAFTK*, Jurkat cells were stimulated with anti-T-cell receptor antibody X35 or OKT-3 or with the lectin PHA. Ligation of the TCR/CD3 by monoclonal antibody X35 or OKT-3 reached a maximum by 2.5-5 min, and declined thereafter. PHA stimulation resulted in an increased tyrosine phosphorylation by 5 min which declined slightly thereafter with substantial phosphorylation still detectable at 20 min. Anti-*RAFTK* immunoblotting of anti-*RAFTK* immunoprecipitates showed that the ~115 Kd phosphotyrosine polypeptide corresponds to the *RAFTK* protein. Depending on the resolution of the gels, *RAFTK* was seen to migrate either as a single band or as a doublet.

***RAFTK* associates with the signaling molecules Fyn and Grb-2.**

To further characterize the role that *RAFTK* plays in T-cell signaling following activation via TCR/CD3 ligation, as well as the other stimulatory pathways activated by the lectin PHA, coimmunoprecipitation studies followed by immunoblotting were performed. A specific association of *RAFTK* with Fyn, a src family kinase which is known to be capable of associating with TCR was observed. A small fraction of Fyn was readily detected as associating with *RAFTK* prior to the TCR/CD3 activation of Jurkat cells and this association increased following stimulation.

The ability of *RAFTK* to form *in vitro* complexes with various SH2 and SH3 domain-containing proteins was then examined. For this purpose, GST-fusion proteins were added to the lysates of the stimulated Jurkat cells, the complexes were immunoprecipitated with glutathione-conjugated beads and the bound proteins were detected by anti-*RAFTK* immunoblotting. Sepharose beads containing the GST-Grb2-SH2 domain and the GST-Fyn-SH2 domain bound *RAFTK* from the activated T-cell lysates. Beads containing only GST, GST-N-terminal Grb2-SH3, GST-C-terminal Grb2-SH3 or GST-Fyn-SH3 failed to bind *RAFTK*, indicating the specificity of these interactions. These data clearly show that the stable interaction between *RAFTK* and the SH2 domain of Fyn can be mimicked *in vitro* and provide additional evidence that the SH2 domain may be the principal determinant of *RAFTK* binding to Fyn *in vivo*.

***RAFTK* associates with the cytoskeletal protein paxillin.**

Following the observation that *RAFTK* may be coimmunoprecipitated with molecules previously characterized as components of the TCR signaling pathways, it was desirable to determine whether certain cytoskeletal molecules in T-cells may also associate with this novel kinase. Using specific antibodies to *RAFTK* or paxillin, a constitutive association of these two molecules was found.

To further investigate the cytoskeletal dependence of the tyrosine phosphorylation of *RAFTK*, Jurkat cells, prior to TCR stimulation, were pre-incubated for 60 min at 37°C with media alone or with cytochalasin D. The phosphorylation of *RAFTK* was reduced following the cytochalasin D treatment of cells.

These studies show that *RAFTK*, a novel signaling molecule that appears to be a member of the FAK family, is present in human T-lymphocytes and participates in signaling pathways following T-cell activation. Following the ligation of the TCR/CD3 there was a robust phosphorylation of *RAFTK* in both the model permanent T-cell lines, Jurkat and H9, as well as in primary PBLs. Parallel studies using other T-cell activators, specifically the lectin PHA, revealed a similar activation of *RAFTK* in a time and concentration dependent manner.

It is noteworthy that following activation, *RAFTK* was found to be associated with several well-characterized components of TCR/CD3 signaling pathways, including Fyn and Grb2. Fyn is known to be capable of associating with the TCR/CD3 complex, and is believed to play an important role in initiating the changes in phosphorylation that lead to further downstream signaling. This role has been most clearly demonstrated in studies showing the impaired development of CD4+CD8+ thymocytes from double mutant mice rendered null for Fyn and FAK through homologous recombination (Kanazawa, S. et al. (1996) *Blood* 87, 865-870). Also, transgenic thymocytes from mice overexpressing Fyn were hyperstimulatable, and overexpression of a catalytically inactive form of Fyn substantially inhibited TCR-mediated activation in otherwise normal thymocytes (Cooke, M. P. et al. (1991) *Cell* 65, 281-291). Grb2 is a well-characterized adaptor molecule that seems capable of associating with a number of kinases and substrates and may also act to facilitate signaling through the enhancement of the physical association of such partners in enzymatic reactions (See, e.g., Li, N. et al. (1993) *Nature* 363, 85-88; Koch, C. A. et al. (1991) *Science* 252, 668-674; Pawson, T., and Gish, G. D. (1992) *Cell* 71, 359-362). Shc and Grb2 have also been shown to play important roles in T-cell signaling (Motto, D. G. et al. (1994) *Journal of Biological Chemistry* 269, 21608-21613; Meisner, H. et al. (1995) *Molecular & Cellular Biology* 15, 3571-3578; Fukazawa, T. et al. (1995) *Journal of Biological Chemistry* 270, 19141-19150).

These observations on *RAFTK* show that this novel signaling molecule can play a variety of roles in the transduction of T-cell signaling. The confluence of signaling molecules and cytoskeletal components provides a platform for the regulated interactions of kinases and substrates and lead to important changes in cell morphology that enable other aspects of the immune response such as adhesion or migration. Analogy with work in adherent mesenchymal cells shows that the formation of the so-called focal adhesions facilitate the creation of these platforms and mediate cell attachment and transduction of signals (See, e.g., Richardson, A. and Parsons, J. T. (1995) *Bioessays* 17, 229; Schaller, M. D. et al. (1992) *Proceedings of the National Academy of Sciences of the United States of America* 89, 5192-5196; Clark, E. A. and Brugge, J. S. (1995) *Science* 268, 233-239). Relatively little is known about similar mechanisms in hematopoietic cells like T-lymphocytes. Recently, another member of the FAK family, termed FAK B, was identified. Initial studies indicated that FAK B may associate with ZAP-70, an intracytoplasmic protein tyrosine kinase also capable of associating with TCR (Kanner, S. B. et al. (1994) *Proceedings of the National Academy of Sciences of the United States of America* 91, 10484-10487).



There is relatively limited information available on the convergence of protein tyrosine kinases and cytoskeletal elements in T-lymphocytes. Several T-cell surface structures, including CD11a/CD18 and CD44, associate with the cytoskeleton upon receptor cross-linking. Recently, the interaction of the chain of TCR with the actin cytoskeleton upon T-cell activation was demonstrated (Valitutti, S. et al. (1995) *Journal of Experimental Medicine* 181, 577-584). These results revealed that *RAFTK* co-associates with paxillin, a major component of the cytoskeleton. Furthermore, the pre-treatment of cells with cytochalasin D results in the reduced tyrosine phosphorylation of *RAFTK* upon T-cell receptor activation. This response shows that *RAFTK* phosphorylation requires the formation of a cytoskeletal complex which provides a foundation for the interactions and compartmentalization of kinases and substrates.

**EXAMPLE 7: Cytokine Signalling Through the Novel Tyrosine Kinase *RAFTK* in Kaposi's Sarcoma Cells**

Considerable data indicate a role of endogenous and exogenous cytokines in the pathogenesis of Kaposi's sarcoma (KS). A number of growth factors including basic FGF, VEGF, oncostatin M (OSM), IL-6, and TNF- $\alpha$  have been reported to promote KS cell growth. A novel tyrosine kinase receptor, FLT-4, was found to be present on normal lymphatic endothelium and robustly expressed in KS cells. Moreover, the recently identified ligand VRP for the FLT-4 receptor results in signalling in KS cells. Signal transduction pathways following receptor engagement by these diverse cytokines that belong to different receptor families was studied. KS cells expressed a recently identified focal adhesion kinase termed *RAFTK* which is believed to coordinate surface signals from cytokine and integrin receptors with the cytoskeleton. *RAFTK* was phosphorylated in KS cells following treatment with b-FGF, OSM, IL-6, VEGF, VRP, or TNF- $\alpha$ . Following *RAFTK* activation by these cytokines, there was enhanced association of *RAFTK* with the cytoskeletal protein paxillin. This association appeared to be mediated through the C-terminal domain of *RAFTK* based on studies using GST-fusion proteins of different *RAFTK* domains. A novel surface receptor FLT-4 expressed on KS cells, as well as a novel intracytoplasmic tyrosine kinase *RAFTK* have been identified. Treatment with diverse cytokines previously reported to potentiate KS cell growth all led to phosphorylation of *RAFTK* and its association with the cytoskeletal protein paxillin. These

observations suggest that inhibition of *RAFTK* may allow for disruption of a common pathway important in KS cell growth and could be clinically exploited as an anti-neoplastic strategy.

5 Kaposi's sarcoma (KS) is the most frequent neoplasm arising among patients with the acquired immune deficiency syndrome (AIDS). The cell of origin of the tumor is believed to be from lymphatic endothelium (Dictor, M. (1988) *Lymphology*. 21, 53-60; Dorfman, R. F. (1988) *Lymphology*. 21, 45-52). Etiological factors implicated in KS include the recently discovered human herpes virus 8 (HHV-8)/Kaposi's sarcoma herpes virus (KSHV) and TAT, the soluble transcriptional  
10 activator of HIV (See, e.g., Chang, Y. et al. (1994) *Science*. 266, 1865-1869; Zhong, W. et al. (1996) *Proc. Natl. Acad. Sci. USA*. 93, 6641-6646; Huang, Y. Q. et al. (1996) *J. Clin. Invest.* 97, 2803-2806). Considerable data indicate a role of endogenous and exogenous cytokines in the pathogenesis of KS (See, e.g., Ba, G. et al. (1992) *J. Immunol.* 149, 3727-3734; Buonaguro, L. et al. (1992) *J. Virol.* 66, 7159-  
15 7167; Cai, J., et al. (1994) *Am. J. Pathol.* 145, 74-79). Growth factors such as basic fibroblast growth factor (b-FGF) and vascular endothelial growth factor (VEGF) which are known to stimulate mitogenesis of certain types of endothelium, as well as oncostatin M, interleukin-6 (IL-6), and tumor necrosis factor alpha (TNF- $\alpha$ ), which are elaborated during inflammatory conditions, have been implicated in promoting KS  
20 cell growth.

Defining the signal transduction pathways which may be utilized by cytokines which appear to stimulate KS growth provides an opportunity for rational and targeted therapeutic intervention against this neoplasm. One issue that is immediately apparent is that the cytokines described to date as promoting KS belong to distinctly  
25 different families as defined by their receptors. VEGF and b-FGF receptors are of the protein tyrosine kinase family, oncostatin M and IL-6 utilize a common gp130 subunit, and TNF- $\alpha$  receptors are of the Fas/apoptosis CD95 family. Cognizant of this diversity, the signalling pathways triggered by cytokine treatment in permanent human KS cells *in vitro* have been characterized and a common molecule sought  
30 among the diverse pathways. KS cells express *RAFTK*.

In KS cells, treatment with cytokines of different families, including b-FGF, oncostatin M, IL-6, VEGF, and TNF- $\alpha$ , all led to phosphorylation of *RAFTK*. After cytokine treatment, *RAFTK* is found to associate with the cytoskeletal protein paxillin. This observation has been extended and focused on the tyrosine kinase receptor  
35 termed FLT-4, which has been found in fetal and adult lymphatic endothelium (Kaipainen A. et al. (1995) *Proc. Natl. Acad. Sci. USA*. 92, 3566-3570; Kaipainen A. et al. (1993) *J. Exp. Med.* 178, 2077-2088; Pajusola K. et al. (1993) *Oncogene*. 8,

2931-2937). KS cells express the FLT-4 receptor, and treatment with its newly discovered ligand called VEGF related protein (VRP or VEGF-C) (Lee, J. et al. (1996) *Proc. Natl. Acad. Sci. USA*. 93, 1988-1992; Joukov V., et al. (1996) *EMBO J.* 15, 290-298) again results in *RAFTK* phosphorylation.

5

The following materials and methods were used to study cytokine signalling through *RAFTK* in Kaposi's Sarcoma cells:

### Cells and cell cultures

10 Human Kaposi's sarcoma cells lines KS 38 were derived from cutaneous biopsy of an AIDS patient as previously described (Lunardi-Iskandar, Y. et al. (1995) *J. Natl. Cancer. Inst.* 87, 974-981; Masood R. et al. (1994) *Human Retroviruses*. 10, 969-975). The cells were grown on 1.5% gelatin-coated flasks and were carried in RPMI 1640 with 15% fetal calf serum (FCS), 2mM glutamine, 1mM MEM Sodium  
15 Pyruvate, 0.05mM MEM Non-Essential Amino Acids, 1x MEM Amino Acids, 1% Nutridoma-HU (Boehringer Mannheim) and 50 mg/mL penicillin and 50 mg/ml streptomycin. Cultures were carried until near confluent prior to different treatments in the signaling studies described below. 293 cells were transfected with the FLT-4 gene and used as controls for detection of receptor protein as described (Lee, J. et al.  
20 (1996) *Proc. Natl. Acad. Sci. USA*. 93, 1988-1992).

### Reagents and antibodies

*RAFTK* antibodies were generated using GST fusion proteins to various domains of the molecule and immunizing New Zealand rabbits as previously  
25 described (Avraham S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751; Li J. et al. (1996) *Blood*. 88, 417-428). Using an ELISA assay, sera were screened for specific binding to *RAFTK*. Serum R4520 was chosen for further studies based on its titer in the ELISA. Serum R4520 did not cross react with FAK and was specific for *RAFTK*. Antibodies to the VEGF receptor FLK-1 and to the receptor FLT-4 were obtained  
30 from Genentech Inc. Antibodies to paxillin were obtained from Santa Cruz Biotechnology. Monoclonal anti-phosphotyrosine antibody (4G10) was obtained from Upstate Biotechnology Inc. (Lake Placid, NY). Electrophoresis reagents were obtained from Bio-Rad Laboratories (Hercules, CA). The Phorbol 12-myristate 13-acetate (PMA) and protease inhibitors leupeptin, aprotinin and alpha 1 antitrypsin and  
35 all other reagents were obtained from Sigma Co. (St.Louis, MO). The recombinant cytokines b-FGF, TNF-a, and IL-6 were obtained from R&D systems. VRP, the ligand for the FLT-4 receptor, was obtained and expressed from a glioblastoma cell

line and purified as previously reported (Lee, J. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93, 1988-1992). Recombinant VEGF was obtained from Genentech, Inc. (South San Francisco, CA). Recombinant oncostatin M was obtained from the AIDS Reagent Bank (Bethesda, MD).

5

### Indirect immunofluorescence

KS 38 cells were cultured in Chamber Slides (Lab Tek) to 90% confluency. Cells were washed twice with cold PBS and then fixed for 30 min in 4% paraformaldehyde. Cells were washed 3X PBS and blocked for non-specific staining using 10% FCS in PBS for 30 minute on ice. FLT-4 and FLK-1 expression were determined using purified antiserum at a dilution 1:100 for one hour on ice. Normal rabbit serum was used as a control for non-specific staining. After washing cells 3X with PBS, cells were stained with secondary antibody conjugated to FITC (Boehringer Mannheim) at 1:500 dilution for one hour on ice. Proteins were visualized and photographed after washing 3X PBS using an inverted fluorescence microscope.

10

15

### Stimulation of cells

KS cells, grown to 80% confluence, were serum-starved for 16-18 hrs and washed twice with Hank's balanced salt solution (Gibco BRL) prior to PMA or cytokine treatments. KS cells were first treated with PMA to assess the effects of a chemical stimulus known to phosphorylate *RAFTK* in other cell systems (Avraham S. et al. (1995) *J. Biol. Chem.* 270, 27742-27751). After a time course of stimulation with PMA was established, the effects of cytokines were studied. VEGF, VRP, TNF- $\alpha$ , oncostatin M, IL-6, or b-FGF were added to cultures at a range of concentrations for different time periods *in vitro*. After stimulation, cell lysates were directly prepared within the culture dish by lysis in 500  $\mu$ L modified RIPA (50 mM Tris-HCl, pH 7.4, 1% NP-40, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM PMSF, 10 (g/ml of aprotinin, leupeptin and pepstatin, 10 mM sodium vanadate, 10 mM sodium fluoride and 10 mM sodium pyrophosphate) per dish at varying timepoints. Total cell lysates (TCL) were clarified by centrifugation at 10,000 x g for 10 min. Protein concentrations were determined by protein assay (Bio-Rad Laboratories).

20

25

30

### Immunoprecipitation and Western blot analysis

For immunoprecipitation studies, identical amounts of protein from each sample were clarified by incubation with protein A-Sepharose CL-4B (Pharmacia Biotech) for 1 h at 4°C. Following the removal of protein A-Sepharose by brief

35

centrifugation, the solution was incubated with different primary antibodies as detailed below for each experiment for 4 h or overnight at 4°C. Immunoprecipitations of the antibody-antigen complexes were performed by incubation for 2 h at 4°C with 75 µl of protein A-Sepharose (10% suspension). Non-specific bound proteins were removed by washing the Sepharose beads three times with modified RIPA buffer and one time with Phosphate buffered saline (PBS). Bound proteins were solubilized in 40 µl of 2 X Laemmli buffer and further analyzed by immunoblotting. Samples were separated on 7.5% SDS-PAGE and then transferred to nitrocellulose membranes. The membranes were blocked with 5% non-fat milk protein and probed with primary antibody for 2 h at RT or 4°C overnight. Immunoreactive bands were visualized using HRP-conjugated secondary antibody and the enhanced chemiluminescent (ECL) system (Amersham Corp., Arlington Heights, IL).

***Kaposi sarcoma cells express Flk-1 and Flt-4 receptors.***

To characterize the effects of different cytokines on KS cell signalling the KS 38 cell line was examined for expression of receptors for members of the VEGF family. Because KS spindle cells appear to be derived from lymphatic endothelium. Using indirect immunofluorescence (IFA), expression of both the FLK-1 receptor and the FLT-4 receptor was readily observed. The presence of both receptors was confirmed by Western blot using specific antisera to FLK-1 and FLT-4.

***RAFTK is expressed in Kaposi sarcoma cells and phosphorylated upon cytokine treatment.***

The KS cells were further characterized for expression of *RAFTK*. KS 38 cells expressed significant amounts of *RAFTK* protein as detected by Western blot and immunoprecipitation. Moreover, PMA treatment of KS 38 cells resulted in a time dependent phosphorylation of *RAFTK*.

Having established that *RAFTK* is expressed in KS 38 cells, whether or not treatment of these cells with VEGF or VRP, respective ligands for the FLK-1 and FLT-4 receptors, resulted in activation of signalling pathways that included *RAFTK* was investigated. There was a clear time dependent phosphorylation of *RAFTK* in response to VRP. Similar changes were observed following treatment with VEGF.

Previously, cytokines such as b-FGF, oncostatin M, IL-6 and TNF-α have been reported to promote *in vitro* proliferation of KS cells. The effects of treatment with these cytokines on *RAFTK* phosphorylation in KS 38 cells was analyzed. Each of these cytokines resulted in phosphorylation of this novel tyrosine kinase.

**Cytokine treatment of Kaposi sarcoma cells results in *RAFTK* association with the cytoskeletal protein paxillin.**

Following the observation that *RAFTK* was phosphorylated by cytokines like oncostatin M, IL-6, b-FGF and TNF- $\alpha$  known to stimulate KS cell growth, as well as the endothelial growth factors VEGF and VRP, whether this phosphorylation might modulate the association of *RAFTK* with certain cytoskeletal molecules was investigated. Using specific antibodies to *RAFTK* or paxillin, significantly increased association of these two molecules following cytokine treatments was found.

The development of authentic permanent KS cell lines has afforded the opportunity to characterize the surface structures of these cells and to examine which cytokines may modulate proliferation of the neoplasm. There is an extensive literature which supports a role for several cytokines in promoting KS cell growth via autocrine or paracrine mechanisms (See, e.g., Ba, G. et al. (1992) *J. Immunol.* 149, 3727-3734; Buonaguro, L. et al. (1992) *J. Virol.* 66, 7159-7167; Cai, J., et al. (1994) *Am. J. Pathol.* 145, 74-79). Characterization of signalling pathways in KS cells and the effects of these cytokines on such pathways have been less extensively explored. Amaral et. al found that OSM activated the MAP kinase pathway (Amaral M. C. et al. (1993) *J. Clin. Invest.* 92, 848-857) while Faris et. al reported that members of the Jak/Stat family of kinases known to participate in signalling via the gp 130 receptor were active in KS cells as well (Faris M. et al. (1996) *AIDS.* 10, 369-378). In these studies, the KS 38 cell line was derived from a patient with cutaneous KS as a model because of its previously characterized properties that closely correspond to those of primary pathological KS specimens (Lunardi-Iskandar, Y. et al. (1995) *J. Natl. Cancer. Inst.* 87, 974-981). The expression of novel receptors on KS 38 cells which are preferentially expressed in normal lymphatic endothelium was investigated, and characterization made of signalling pathways that may link surface receptor activation to the cytoskeleton in these cells.

The tyrosine kinase FLT-4 receptor is relatively restricted in expression in normal tissues, with prior studies indicating its presence on the surface of lymphatic endothelium (Kaipainen A. et al. (1995) *Proc. Natl. Acad. Sci. USA.* 92, 3566-3570; Kaipainen A. et al. (1993) *J. Exp. Med.* 178, 2077-2088; Pajusola K. et al. (1993) *Oncogene.* 8, 2931-2937). KS 38 cells express FLT-4 as well as the related FLK-1 receptor. The recently identified ligand VRP specifically binds to the FLT-4 receptor (Lee, J. et al. (1996) *Proc. Natl. Acad. Sci. USA.* 93, 1988-1992; Joukov V., et al. (1996) *EMBO J.* 15, 290-298), while FLK-1 is activated by VEGF. VRP, as well as VEGF, induced significant signalling changes in target KS 38 cells based on enhanced phosphorylation of proteins. Following this observation, investigation into signalling

molecules whose enhanced phosphorylation was a common response to these cytokines as well as those previously reported to stimulate KS cells was performed.

A variety of ligands and receptors of different molecular families have been implicated in the pathogenesis of KS. To survey this range of cytokines,  
5 representative cytokines from each family was chosen and a comparative analysis was made of b-FGF, TNF- $\alpha$ , oncostatin M, and IL-6 with VEGF and VRP. They all signalled via the recently identified *RAFTK* molecule.

*RAFTK* functions as a "platform kinase" upon which a number of intracytoplasmic kinases and adaptor molecules converge. The convergence of such  
10 molecules facilitates transmission of surface signals to the cytoskeleton. In this study, the phosphorylation of *RAFTK* by a variety of cytokines which belong to distinctly different families is described. Specifically, b-FGF, VEGF and VRP signal through receptor tyrosine kinases, OSM and IL-6 bind to a dual receptor with a specific chain and a shared gp130 chain, and TNF- $\alpha$  binds to the CD95 receptor family linked to  
15 apoptosis. It appears, from this KS cell line model, that *RAFTK* may participate in each of these diverse receptor activated pathways. Based on the observations reported here, *RAFTK* functions in KS cells to transduce receptor signals via association with cytoskeletal molecules such as paxillin. *RAFTK* activation likely participates in a final common pathway for KS cell growth. Given the accessibility of cutaneous KS  
20 lesions to locally applied treatments, specific inhibitors of *RAFTK* are particularly useful in treatment of this disorder.

### Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more  
25 than routine experimentation, many equivalents of the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- 5 (i) APPLICANTS: Shalom Avraham et al.
- (ii) TITLE OF INVENTION: Novel RAFTK Signaling Molecules and Uses Therefor
- 10 (iii) NUMBER OF SEQUENCES: 18
- (iv) CORRESPONDENCE ADDRESS:
- 15 (A) ADDRESSEE: LAHIVE & COCKFIELD  
(B) STREET: 60 State Street, suite 510  
(C) CITY: Boston  
(D) STATE: Massachusetts  
(E) COUNTRY: USA  
(F) ZIP: 02109-1875
- 20 (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
25 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:  
(B) FILING DATE: August 23, 1996  
30 (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER:  
(B) FILING DATE:
- 35 (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Jean M. Silveri  
(B) REGISTRATION NUMBER: 39,030  
(C) REFERENCE/DOCKET NUMBER: NER-255
- 40 (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: (617)227-7400  
(B) TELEFAX: (617)227-5941
- 45 (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3621 base pairs  
(B) TYPE: nucleic acid  
50 (C) STRANDEDNESS: single  
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- 55 (ix) FEATURE:



(A) NAME/KEY: CDS

(B) LOCATION: 294..3321

5	(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:	
	AACCACAAGT CAAATAGAAA GAAGTTAAAA GAATGTTTAT GCAAACACAT GAGAAAAGAA	60
10	GGGTGCAGAT GAGAATAGGG GTGTGGTTAA CAACTCAGAG GAGGAGGGAG AATCTAACCT	120
	GTCAGCCCTT TTA CT CAGCC ACAGCCTCCG GAGCCGTTGC ACACCTACCT GCCCGGCCGA	180
	CTTACCTGTA CTTGCCGCCG TCCCGGCTCA CCTGGCGGTG CCCGAGGAGT AGTCGCTGGA	240
15	GTCCGCGCCT CCCTGGGACT GCAATGTGCC GATCTTAGCT GCTGCCTGAG AGG ATG	296
	Met	
	1	
20	TCT GGG GTG TCC GAG CCC CTG AGT CGA GTA AAG TTG GGC ACG TTA CGC	344
	Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Leu Gly Thr Leu Arg	
	5 10 15	
25	CGG CCT GAA GGC CCT GCA GAG CCC ATG GTG GTG GTA CCA GTA GAT GTG	392
	Arg Pro Glu Gly Pro Ala Glu Pro Met Val Val Val Pro Val Asp Val	
	20 25 30	
30	GAA AAG GAG GAC GTG CGT ATC CTC AAG GTC TGC TTC TAT AGC AAC AGC	440
	Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn Ser	
	35 40 45	
35	TTC AAT CCT GGG AAA AAC TTC AAA CTG GTC AAA TGC ACT GTC CAG ACG	488
	Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln Thr	
	50 55 60 65	
40	GAG ATC CGG GAG ATC ATC ACC TCC ATC CTG CTG AGC GGG CGG ATC GGG	536
	Glu Ile Arg Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile Gly	
	70 75 80	
45	CCC AAC ATC CGG TTG GCT GAG TGC TAT GGG CTG AGG CTG AAG CAC ATG	584
	Pro Asn Ile Arg Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His Met	
	85 90 95	
50	AAG TCC GAT GAG ATC CAC TGG CTG CAC CCA CAG ATG ACG GTG GGT GAG	632
	Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly Glu	
	100 105 110	
55	GTG CAG GAC AAG TAT GAG TGT CTG CAC GTG GAA GCC GAG TGG AGG TAT	680
	Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg Tyr	
	115 120 125	
60	GAC CTT CAA ATC CGC TAC TTG CCA GAA GAC TTC ATG GAG AGC CTG AAG	728
	Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu Lys	
	130 135 140 145	
65	GAG GAC AGG ACC ACG CTG CTC TAT TTT TAC CAA CAG CTC CGG AAC GAC	776
	Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn Asp	

- 132 -

					150					155					160						
					TAC	GCC	AGC	AAG	GTC	AGC	GAG	GGC	ATG	GCC	CTG	CAG			824		
5					Tyr	Ala	Ser	Lys	Val	Ser	Glu	Gly	Met	Ala	Leu	Gln					
					165				170					175							
					CTG	GAG	CTC	AGG	CGG	TTC	TTC	AAG	GAT	ATG	CCC	CAC	AAT		872		
					Leu	Glu	Leu	Arg	Arg	Phe	Phe	Lys	Asp	Met	Pro	His	Asn				
10					180				185					190							
					GCA	CTT	GAC	AAG	AAG	TCC	AAC	TTC	GAG	CTC	CTA	GAA	AAG	GAA	GTG	GGG	920
					Ala	Leu	Asp	Lys	Lys	Ser	Asn	Phe	Glu	Leu	Leu	Glu	Lys	Glu	Val	Gly	
					195				200						205						
15					CTG	GAC	TTG	TTT	TTC	CCA	AAG	CAG	ATG	CAG	GAG	AAC	TTA	AAG	CCC	AAA	968
					Leu	Asp	Leu	Phe	Phe	Pro	Lys	Gln	Met	Gln	Glu	Asn	Leu	Lys	Pro	Lys	
					210				215						220				225		
					CAG	TTC	CGG	AAG	ATG	ATC	CAG	CAG	ACC	TTC	CAG	CAG	TAC	GCC	TCG	CTC	1016
20					Gln	Phe	Arg	Lys	Met	Ile	Gln	Gln	Thr	Phe	Gln	Gln	Tyr	Ala	Ser	Leu	
					230									235					240		
					AGG	GAG	GAG	GAG	TGC	GTC	ATG	AAG	TTC	TTC	AAC	ACT	CTC	GCC	GGC	TTC	1064
25					Arg	Glu	Glu	Glu	Cys	Val	Met	Lys	Phe	Phe	Asn	Thr	Leu	Ala	Gly	Phe	
					245								250					255			
					GCC	AAC	ATC	GAC	CAG	GAG	ACC	TAC	CGC	TGT	GAA	CTC	ATT	CAA	GGA	TGG	1112
					Ala	Asn	Ile	Asp	Gln	Glu	Thr	Tyr	Arg	Cys	Glu	Leu	Ile	Gln	Gly	Trp	
					260							265						270			
30					AAC	ATT	ACT	GTG	GAC	CTG	GTC	ATT	GGC	CCT	AAA	GGG	ATC	CGC	CAG	CTG	1160
					Asn	Ile	Thr	Val	Asp	Leu	Val	Ile	Gly	Pro	Lys	Gly	Ile	Arg	Gln	Leu	
					275							280					285				
35					ACT	AGT	CAG	GAC	GCA	AAG	CCC	ACC	TGC	CTG	GCC	GAG	TTC	AAG	CAG	ATC	1208
					Thr	Ser	Gln	Asp	Ala	Lys	Pro	Thr	Cys	Leu	Ala	Glu	Phe	Lys	Gln	Ile	
					290					295					300					305	
					AGG	TCC	ATC	AGG	TGC	CTC	CCG	CTG	GAG	GAG	GGC	CAG	GCA	GTA	CTT	CAG	1256
40					Arg	Ser	Ile	Arg	Cys	Leu	Pro	Leu	Glu	Glu	Gly	Gln	Ala	Val	Leu	Gln	
					310									315					320		
					CTG	GGC	ATT	GAA	GGT	GCC	CCC	CAG	GCC	TTG	TCC	ATC	AAA	ACC	TCA	TCC	1304
45					Leu	Gly	Ile	Glu	Gly	Ala	Pro	Gln	Ala	Leu	Ser	Ile	Lys	Thr	Ser	Ser	
					325								330					335			
					CTA	GCA	GAG	GCT	GAG	AAC	ATG	GCT	GAC	CTC	ATA	GAC	GGC	TAC	TGC	CGG	1352
					Leu	Ala	Glu	Ala	Glu	Asn	Met	Ala	Asp	Leu	Ile	Asp	Gly	Tyr	Cys	Arg	
					340							345					350				
50					CTG	CAG	GGT	GAG	CAC	CAA	GGC	TCT	CTC	ATC	ATC	CAT	CCT	AGG	AAA	GAT	1400
					Leu	Gln	Gly	Glu	His	Gln	Gly	Ser	Leu	Ile	Ile	His	Pro	Arg	Lys	Asp	
					355						360						365				
55					GGT	GAG	AAG	CGG	AAC	AGC	CTG	CCC	CAG	ATC	CCC	ATG	CTA	AAC	CTG	GAG	1448
					Gly	Glu	Lys	Arg	Asn	Ser	Leu	Pro	Gln	Ile	Pro	Met	Leu	Asn	Leu	Glu	

	370	375	380	385	
5	GCC CGG CGG TCC CAC CTC TCA GAG AGC TGC AGC ATA GAG TCA GAC ATC Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp Ile	390	395	400	1496
10	TAC GCA GAG ATT CCC GAC GAA ACC CTG CGA AGG CCC GGA GGT CCA CAG Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro Gln	405	410	415	1544
15	TAT GGC ATT GCC CGT GAA GAT GTG GTC CTG AAT CGT ATT CTT GGG GAA Tyr Gly Ile Ala Arg Glu Asp Val Val Leu Asn Arg Ile Leu Gly Glu	420	425	430	1592
20	GGC TTT TTT GGG GAG GTC TAT GAA GGT GTC TAC ACA AAT CAT AAA GGG Gly Phe Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys Gly	435	440	445	1640
25	GAG AAA ATC AAT GTA GCT GTC AAG ACC TGC AAG AAA GAC TGC ACT CTG Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr Leu	450	455	460	1688
30	GAC AAC AAG GAG AAG TTC ATG AGC GAG GCA GTG ATC ATG AAG AAC CTC Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn Leu	470	475	480	1736
35	GAC CAC CCG CAC ATC GTG AAG CTG ATC GGC ATC ATT GAA GAG GAG CCC Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu Pro	485	490	495	1784
40	ACC TGG ATC ATC ATG GAA TTG TAT CCC TAT GGG GAG CTG GGC CAC TAC Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His Tyr	500	505	510	1832
45	CTG GAG CGG AAC AAG AAC TCC CTG AAG GTG CTC ACC CTC GTG CTG TAC Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Leu Thr Leu Val Leu Tyr	515	520	525	1880
50	TCA CTG CAG ATA TGC AAA GCC ATG GCC TAC CTG GAG AGC ATC AAC TGC Ser Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn Cys	530	535	540	1928
55	GTG CAC AGG GAC ATT GCT GTC CGG AAC ATC CTG GTG GCC TCC CCT GAG Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro Glu	550	555	560	1976
60	TGT GTG AAG CTG GGG GAC TTT GGT CTT TCC CGG TAC ATT GAG GAC GAG Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp Glu	565	570	575	2024
65	GAC TAT TAC AAA GCC TCT GTG ACT CGT CTC CCC ATC AAA TGG ATG TCC Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met Ser	580	585	590	2072
70	CCA GAG TCC ATT AAC TTC CGA CGC TTC ACG ACA GCC AGT GAC GTC TGG Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val Trp				2120

	595	600	605	
5	ATG TTC GCC GTG TGC ATG TGG GAG ATC CTG AGC TTT GGG AAG CAG CCC Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln Pro 610 615 620 625	2168		
10	TTC TTC TGG CTG GAG AAC AAG GAT GTC ATC GGG GTG CTG GAG AAA GGA Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys Gly 630 635 640	2216		
15	GAC CGG CTG CCC AAG CCT GAT CTC TGT CCA CCG GTC CTT TAT ACC CTC Asp Arg Leu Pro Lys Pro Asp Leu Cys Pro Pro Val Leu Tyr Thr Leu 645 650 655	2264		
20	ATG ACC CGC TGC TGG GAC TAC GAC CCC AGT GAC CGG CCC CGC TTC ACC Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe Thr 660 665 670	2312		
25	GAG CTG GTG TGC AGC CTC AGT GAC GTT TAT CAG ATG GAG AAG GAC ATT Glu Leu Val Cys Ser Leu Ser Asp Val Tyr Gln Met Glu Lys Asp Ile 675 680 685	2360		
30	GCC ATG GAG CAA GAG AGG AAT GCT CGC TAC CGA ACC CCC AAA ATC TTG Ala Met Glu Gln Glu Arg Asn Ala Arg Tyr Arg Thr Pro Lys Ile Leu 690 695 700 705	2408		
35	GAG CCC ACA GCC TTC CAG GAA CCC CCA CCC AAG CCC AGC CGA CCT AAG Glu Pro Thr Ala Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro Lys 710 715 720	2456		
40	TAC AGA CCC CCT CCG CAA ACC AAC CTC CTG GCT CCA AAG CTG CAG TTC Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln Phe 725 730 735	2504		
45	CAG GTT CCT GAG GGT CTG TGT GCC AGC TCT CCT ACG CTC ACC AGC CCT Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser Pro 740 745 750	2552		
50	ATG GAG TAT CCA TCT CCC GTT AAC TCA CTG CAC ACC CCA CCT CTC CAC Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu His 755 760 765	2600		
55	CGG CAC AAT GTC TTC AAA CGC CAC AGC ATG CGG GAG GAG GAC TTC ATC Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe Ile 770 775 780 785	2648		
60	CAA CCC AGC AGC CGA GAA GAG GCC CAG CAG CTG TGG GAG GCT GAA AAG Gln Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu Lys 790 795 800	2696		
65	GTC AAA ATG CGG CAA ATC CTG GAC AAA CAG CAG AAG CAG ATG GTG GAG Val Lys Met Arg Gln Ile Leu Asp Lys Gln Gln Lys Gln Met Val Glu 805 810 815	2744		
70	GAC TAC CAG TGG CTC AGG CAG GAG GAG AAG TCC CTG GAC CCC ATG GTT Asp Tyr Gln Trp Leu Arg Gln Glu Glu Lys Ser Leu Asp Pro Met Val	2792		

	820	825	830	
5	TAT ATG AAT GAT AAG TCC CCA TTG ACG CCA GAG AAG GAG GTC GGC TAC Tyr Met Asn Asp Lys Ser Pro Leu Thr Pro Glu Lys Glu Val Gly Tyr 835 840 845			2840
10	CTG GAG TTC ACA GGG CCC CCA CAG AAG CCC CCG AGG CTG GGC GCA CAG Leu Glu Phe Thr Gly Pro Pro Gln Lys Pro Pro Arg Leu Gly Ala Gln 850 855 860 865			2888
15	TCC ATC CAG CCC ACA GCT AAC CTG GAC CGG ACC GAT GAC CTG GTG TAC Ser Ile Gln Pro Thr Ala Asn Leu Asp Arg Thr Asp Asp Leu Val Tyr 870 875 880			2936
20	CTC AAT GTC ATG GAG CTG GTG CGG GCC GTG CTG GAG CTC AAG AAT GAG Leu Asn Val Met Glu Leu Val Arg Ala Val Leu Glu Leu Lys Asn Glu 885 890 895			2984
25	CTC TGT CAG CTG CCC CCC GAG GGC TAC GTG GTG GTG GTG AAG AAT GTG Leu Cys Gln Leu Pro Pro Glu Gly Tyr Val Val Val Val Lys Asn Val 900 905 910			3032
30	GGG CTG ACC CTG CGG AAG CTC ATC GGG AGC GTG GAT GAT CTC CTG CCT Gly Leu Thr Leu Arg Lys Leu Ile Gly Ser Val Asp Asp Leu Leu Pro 915 920 925			3080
35	TCC TTG CCG TCA TCT TCA CGG ACA GAG ATC GAG GGC ACC CAG AAA CTG Ser Leu Pro Ser Ser Ser Arg Thr Glu Ile Glu Gly Thr Gln Lys Leu 930 935 940 945			3128
40	CTC AAC AAA GAC CTG GCA GAG CTC ATC AAC AAG ATG CGG CTG GCG CAG Leu Asn Lys Asp Leu Ala Glu Leu Ile Asn Lys Met Arg Leu Ala Gln 950 955 960			3176
45	CAG AAC GCC GTG ACC TCC CTG AGT GAG GAG TGC AAG AGG CAG ATG CTG Gln Asn Ala Val Thr Ser Leu Ser Glu Glu Cys Lys Arg Gln Met Leu 965 970 975			3224
50	ACG GCT TCA CAC ACC CTG GCT GTG GAC GCC AAG AAC CTG CTC GAC GCT Thr Ala Ser His Thr Leu Ala Val Asp Ala Lys Asn Leu Leu Asp Ala 980 985 990			3272
55	GTG GAC CAG GCC AAG GTT CTG GCC AAT CTG GCC CAC CCA CCT GCA GAG T Val Asp Gln Ala Lys Val Leu Ala Asn Leu Ala His Pro Pro Ala Glu 995 1000 1005			3321
	GACGGAGGGT GGGGGCCACC TGCCTGCGTC TTCCGCCCTT GCCTGCCATG TACCTCCCCT			3381
	GCCTTGCTGT TGGTCATGTG GGTCTTCCAG GGAGAAGGCC AAGGGGAGTC ACCTTCCCTT			3441
	GCCACTTTGC ACGACGCCCT CTCCCCACCC CTACCCCTGG CTGTACTGCT CAGGCTGCAG			3501
	CTGGACAGAG GGGACTCTGG GCTATGGACA CAGGGTGACG GTGACAAAGA TGGCTCAGAG			3561
	GGGGACTGCT GCTGCCTGGC CACTGCTCCC TAAGCCAGCC TGGTCCATGC AGGGGGCTCG			3621

## (2) INFORMATION FOR SEQ ID NO:2:

- 5 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1009 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

- 10 (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

15 Met Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Leu Gly Thr Leu  
 1 5 10 15  
 Arg Arg Pro Glu Gly Pro Ala Glu Pro Met Val Val Val Pro Val Asp  
 20 20 25 30  
 20 Val Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn  
 35 40 45  
 Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln  
 50 55 60  
 25 Thr Glu Ile Arg Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile  
 65 70 75 80  
 Gly Pro Asn Ile Arg Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His  
 30 85 90 95  
 Met Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly  
 100 105 110  
 35 Glu Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg  
 115 120 125  
 Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu  
 130 135 140  
 40 Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn  
 145 150 155 160  
 Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val Ser Glu Gly Met Ala Leu  
 45 165 170 175  
 Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe Phe Lys Asp Met Pro His  
 180 185 190  
 50 Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu Leu Leu Glu Lys Glu Val  
 195 200 205  
 Gly Leu Asp Leu Phe Phe Pro Lys Gln Met Gln Glu Asn Leu Lys Pro  
 210 215 220  
 55 Lys Gln Phe Arg Lys Met Ile Gln Gln Thr Phe Gln Gln Tyr Ala Ser

- 137 -

	225		230		235		240
	Leu Arg Glu Glu Glu Cys Val Met Lys Phe Phe Asn Thr Leu Ala Gly						
		245		250		255	
5	Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg Cys Glu Leu Ile Gln Gly						
		260		265		270	
10	Trp Asn Ile Thr Val Asp Leu Val Ile Gly Pro Lys Gly Ile Arg Gln						
		275		280		285	
	Leu Thr Ser Gln Asp Ala Lys Pro Thr Cys Leu Ala Glu Phe Lys Gln						
		290		295		300	
15	Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu Glu Gly Gln Ala Val Leu						
		305		310		315	320
	Gln Leu Gly Ile Glu Gly Ala Pro Gln Ala Leu Ser Ile Lys Thr Ser						
		325		330		335	
20	Ser Leu Ala Glu Ala Glu Asn Met Ala Asp Leu Ile Asp Gly Tyr Cys						
		340		345		350	
25	Arg Leu Gln Gly Glu His Gln Gly Ser Leu Ile Ile His Pro Arg Lys						
		355		360		365	
	Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln Ile Pro Met Leu Asn Leu						
		370		375		380	
30	Glu Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp						
		385		390		395	400
	Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro						
		405		410		415	
35	Gln Tyr Gly Ile Ala Arg Glu Asp Val Val Leu Asn Arg Ile Leu Gly						
		420		425		430	
40	Glu Gly Phe Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys						
		435		440		445	
	Gly Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr						
		450		455		460	
45	Leu Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn						
		465		470		475	480
	Leu Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu						
		485		490		495	
50	Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His						
		500		505		510	
55	Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Leu Thr Leu Val Leu						
		515		520		525	

- 138 -

Tyr Ser Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn  
 530 535 540

5 Cys Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro  
 545 550 555 560

Glu Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp  
 565 570 575

10 Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met  
 580 585 590

Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val  
 595 600 605

15 Trp Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln  
 610 615 620

20 Pro Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys  
 625 630 635 640

Gly Asp Arg Leu Pro Lys Pro Asp Leu Cys Pro Pro Val Leu Tyr Thr  
 645 650 655

25 Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe  
 660 665 670

Thr Glu Leu Val Cys Ser Leu Ser Asp Val Tyr Gln Met Glu Lys Asp  
 675 680 685

30 Ile Ala Met Glu Gln Glu Arg Asn Ala Arg Tyr Arg Thr Pro Lys Ile  
 690 695 700

35 Leu Glu Pro Thr Ala Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro  
 705 710 715 720

Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln  
 725 730 735

40 Phe Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser  
 740 745 750

Pro Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu  
 755 760 765

45 His Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe  
 770 775 780

50 Ile Gln Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu  
 785 790 795 800

Lys Val Lys Met Arg Gln Ile Leu Asp Lys Gln Gln Lys Gln Met Val  
 805 810 815

55 Glu Asp Tyr Gln Trp Leu Arg Gln Glu Glu Lys Ser Leu Asp Pro Met  
 820 825 830



Val Tyr Met Asn Asp Lys Ser Pro Leu Thr Pro Glu Lys Glu Val Gly  
835 840 845

5 Tyr Leu Glu Phe Thr Gly Pro Pro Gln Lys Pro Pro Arg Leu Gly Ala  
850 855 860

Gln Ser Ile Gln Pro Thr Ala Asn Leu Asp Arg Thr Asp Asp Leu Val  
865 870 875 880

10 Tyr Leu Asn Val Met Glu Leu Val Arg Ala Val Leu Glu Leu Lys Asn  
885 890 895

Glu Leu Cys Gln Leu Pro Pro Glu Gly Tyr Val Val Val Val Lys Asn  
900 905 910

Val Gly Leu Thr Leu Arg Lys Leu Ile Gly Ser Val Asp Asp Leu Leu  
915 920 925

20 Pro Ser Leu Pro Ser Ser Ser Arg Thr Glu Ile Glu Gly Thr Gln Lys  
930 935 940

Leu Leu Asn Lys Asp Leu Ala Glu Leu Ile Asn Lys Met Arg Leu Ala  
945 950 955 960

25 Gln Gln Asn Ala Val Thr Ser Leu Ser Glu Glu Cys Lys Arg Gln Met  
965 970 975

Leu Thr Ala Ser His Thr Leu Ala Val Asp Ala Lys Asn Leu Leu Asp  
980 985 990

Ala Val Asp Gln Ala Lys Val Leu Ala Asn Leu Ala His Pro Pro Ala  
995 1000 1005

35 Glu

- 40 (2) INFORMATION FOR SEQ ID NO:3:
- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4029 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- 50 (ix) FEATURE:
- (A) NAME/KEY: CDS
  - (B) LOCATION: 202..3229
- 55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

	GAGAGCAGCA GGGGTGTGGT TAACGACCGA GAGGAGGAGG GGGAAAAACA ACCTGTCAGC	60
	CTCTTACTCA GCCTCTGCAG GCAGAGCCGC GCGTCCTACC TGCGGCGGCT GCGCTCACCT	120
5	GGCCCAGCCC GGAGCCCTGG CCCGAGTCCG CGCCTCGCCC GAGGGACTGC AATGTGCCGG	180
	TCCTAGCTGC AGTCTGAGAG G ATG TCC GGG GTG TCT GAG CCC TTG AGC CGT	231
	Met Ser Gly Val Ser Glu Pro Leu Ser Arg	
	1 5 10	
10	GTA AAA GTG GGC ACT TTA CGC CGG CCT GAG GGC CCC CCA GAG CCC ATG	279
	Val Lys Val Gly Thr Leu Arg Arg Pro Glu Gly Pro Pro Glu Pro Met	
	15 20 25	
15	GTG GTG GTA CCA GTG GAT GTG GAG AAG GAA GAC GTG CGC ATC CTC AAG	327
	Val Val Val Pro Val Asp Val Glu Lys Glu Asp Val Arg Ile Leu Lys	
	30 35 40	
20	GTC TGC TTC TAC AGC AAC AGC TTC AAC CCA GGG AAG AAC TTC AAG CTT	375
	Val Cys Phe Tyr Ser Asn Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu	
	45 50 55	
25	GTC AAA TGC ACA GTG CAG ACA GAG ATC CAG GAG ATC ATC ACC TCC ATC	423
	Val Lys Cys Thr Val Gln Thr Glu Ile Gln Glu Ile Ile Thr Ser Ile	
	60 65 70	
30	CTC CTG AGT GGG CGA ATA GGG CCC AAC ATC CAG CTG GCT GAA TGC TAT	471
	Leu Leu Ser Gly Arg Ile Gly Pro Asn Ile Gln Leu Ala Glu Cys Tyr	
	75 80 85 90	
35	GGG CTG AGG CTG AAG CAC ATG AAG TCA GAC GAG ATC CAC TGG CTG CAC	519
	Gly Leu Arg Leu Lys His Met Lys Ser Asp Glu Ile His Trp Leu His	
	95 100 105	
40	CCA CAG ATG ACC GTG GGC GAA GTG CAG GAC AAG TAT GAA TGT CTA CAC	567
	Pro Gln Met Thr Val Gly Glu Val Gln Asp Lys Tyr Glu Cys Leu His	
	110 115 120	
45	GTG GAA GCT GAG TGG AGG TAT GAC CTT CAA ATC CGC TAC TTG CCG GAA	615
	Val Glu Ala Glu Trp Arg Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu	
	125 130 135	
50	GAC TTC ATG GAG AGC CTG AAA GAA GAC AGG ACC ACA TTG CTG TAC TTT	663
	Asp Phe Met Glu Ser Leu Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe	
	140 145 150	
55	TAT CAA CAG CTC CGG AAT GAC TAC ATG CAA CGC TAC GCC AGC AAG GTC	711
	Tyr Gln Gln Leu Arg Asn Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val	
	155 160 165 170	
55	AGT GAA GGC ATG GCT CTG CAG CTG GGC TGT CTG GAG CTC AGG AGA TTC	759
	Ser Glu Gly Met Ala Leu Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe	
	175 180 185	
55	TTC AAG GAC ATG CCC CAC AAT GCA CTG GAC AAA AAG TCC AAC TTT GAA	807
	Phe Lys Asp Met Pro His Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu	

	190	195	200	
5	CTC CTG GAA AAA GAA GTC GGT CTG GAC CTG TTT TTC CCA AAG CAG ATG Leu Leu Glu Lys Glu Val Gly Leu Asp Leu Phe Phe Pro Lys Gln Met	205 210	215	855
10	CAG GAA AAC TTA AAG CCC AAG CAG TTC CGG AAG ATG ATC CAG CAG ACC Gln Glu Asn Leu Lys Pro Lys Gln Phe Arg Lys Met Ile Gln Gln Thr	220 225	230	903
15	TTC CAG CAG TAT GCA TCA CTC CGG GAG GAA GAG TGT GTC ATG AAA TTC Phe Gln Gln Tyr Ala Ser Leu Arg Glu Glu Glu Cys Val Met Lys Phe	235 240	245 250	951
20	TTC AAT ACC CTA GCG GGC TTT GCC AAC ATT GAC CAG GAG ACC TAC CGC Phe Asn Thr Leu Ala Gly Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg	255 260	265	999
25	TGC GAA CTC ATT CAA GGA TGG AAC ATT ACT GTG GAC CTG GTC ATC GGC Cys Glu Leu Ile Gln Gly Trp Asn Ile Thr Val Asp Leu Val Ile Gly	270 275	280	1047
30	CCT AAA GGC ATC CGT CAG CTG ACA AGT CAA GAT ACA AAG CCC ACC TGC Pro Lys Gly Ile Arg Gln Leu Thr Ser Gln Asp Thr Lys Pro Thr Cys	285 290	295	1095
35	CTG GCC GAG TTT AAG CAG ATC AGA TCC ATC AGG TGC CTC CCA TTG GAA Leu Ala Glu Phe Lys Gln Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu	300 305	310	1143
40	GAG ACC CAG GCA GTC CTG CAG CTG GGC ATC GAG GGT GCC CCC CAG TCC Glu Thr Gln Ala Val Leu Gln Leu Gly Ile Glu Gly Ala Pro Gln Ser	315 320	325 330	1191
45	TTG TCT ATC AAA ACG TCG TCC CTG GCA GAG GCT GAG AAC ATG GCT GAT Leu Ser Ile Lys Thr Ser Ser Leu Ala Glu Ala Glu Asn Met Ala Asp	335 340	345	1239
50	CTC ATA GAT GGC TAC TGC AGG CTG CAA GGA GAA CAT AAG GGC TCT CTC Leu Ile Asp Gly Tyr Cys Arg Leu Gln Gly Glu His Lys Gly Ser Leu	350 355	360	1287
55	ATC ATG CAT GCC AAG AAA GAT GGT GAG AAG AGG AAC AGC CTG CCT CAG Ile Met His Ala Lys Lys Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln	365 370	375	1335
60	ATC CCC ACA CTA AAC CTG GAG GCT CGG CGG TCG CAC CTC TCA GAA AGC Ile Pro Thr Leu Asn Leu Glu Ala Arg Arg Ser His Leu Ser Glu Ser	380 385	390	1383
65	TGC AGC ATA GAG TCA GAC ATC TAT GCG GAG ATT CCC GAT GAG ACC CTG Cys Ser Ile Glu Ser Asp Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu	395 400	405 410	1431
70	CGA AGA CCA GGA GGT CCA CAG TAC GGT GTT GCC CGT GAA GAA GTA GTT Arg Arg Pro Gly Gly Pro Gln Tyr Gly Val Ala Arg Glu Glu Val Val			1479

- 142 -

	415	420	425	
5	CTT AAC CGC ATT CTG GGT GAA GGC TTC TTT GGG GAG GTC TAT GAA GGT Leu Asn Arg Ile Leu Gly Glu Gly Phe Phe Gly Glu Val Tyr Glu Gly 430 435 440	1527		
10	GTC TAC ACG AAC CAC AAA GGG GAA AAA ATT AAT GTG GCC GTC AAG ACC Val Tyr Thr Asn His Lys Gly Glu Lys Ile Asn Val Ala Val Lys Thr 445 450 455	1575		
15	TGT AAG AAA GAC TGT ACC CAG GAC AAC AAG GAG AAG TTC ATG AGT GAG Cys Lys Lys Asp Cys Thr Gln Asp Asn Lys Glu Lys Phe Met Ser Glu 460 465 470	1623		
20	GCA GTG ATC ATG AAG AAT CTT GAC CAC CCT CAC ATC GTG AAG CTG ATT Ala Val Ile Met Lys Asn Leu Asp His Pro His Ile Val Lys Leu Ile 475 480 485 490	1671		
25	GGC ATC ATT GAA GAG GAA CCC ACC TGG ATT ATC ATG GAA CTG TAT CCT Gly Ile Ile Glu Glu Glu Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro 495 500 505	1719		
30	TAT GGG GAG CTG GGA CAC TAC CTG GAA CGA AAT AAA AAC TCC CTG AAG Tyr Gly Glu Leu Gly His Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys 510 515 520	1767		
35	GTA CCC ACT CTG GTC CTG TAC ACC CTA CAG ATA TGC AAA GCC ATG GCC Val Pro Thr Leu Val Leu Tyr Thr Leu Gln Ile Cys Lys Ala Met Ala 525 530 535	1815		
40	TAT CTG GAG AGC ATC AAC TGT GTG CAC AGG GAT ATT GCT GTC CGG AAC Tyr Leu Glu Ser Ile Asn Cys Val His Arg Asp Ile Ala Val Arg Asn 540 545 550	1863		
45	ATC CTG GTG GCC TCT CCT GAG TGT GTG AAG CTG GGG GAC TTT GGG CTC Ile Leu Val Ala Ser Pro Glu Cys Val Lys Leu Gly Asp Phe Gly Leu 555 560 565 570	1911		
50	TCC CGG TAC ATT GAG GAC GAA GAC TAT TAC AAA GCC TCT GTG ACA CGT Ser Arg Tyr Ile Glu Asp Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg 575 580 585	1959		
55	CTA CCC ATC AAA TGG ATG TCC CCC GAG TCC ATC AAC TTC CGC CGC TTC Leu Pro Ile Lys Trp Met Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe 590 595 600	2007		
60	ACA ACC GCC AGT GAT GTC TGG ATG TTT GCT GTA TGC ATG TGG GAG ATC Thr Thr Ala Ser Asp Val Trp Met Phe Ala Val Cys Met Trp Glu Ile 605 610 615	2055		
65	CTC AGC TTT GGG AAG CAG CCT TTC TTC TGG CTC GAA AAT AAG GAT GTC Leu Ser Phe Gly Lys Gln Pro Phe Phe Trp Leu Glu Asn Lys Asp Val 620 625 630	2103		
70	ATC GGA GTG CTG GAG AAA GGG GAC AGG CTG CCC AAG CCC GAA CTC TGT Ile Gly Val Leu Glu Lys Gly Asp Arg Leu Pro Lys Pro Glu Leu Cys 635 640 645 650	2151		

	635	640	645	650	
5	CCG CCT GTC CTT TAC ACA CTC ATG ACT CGC TGC TGG GAC TAC GAC CCC Pro Pro Val Leu Tyr Thr Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro	655	660	665	2199
10	AGT GAC CGG CCC CGC TTC ACG GAG CTT GTG TGC AGC CTC AGT GAC ATT Ser Asp Arg Pro Arg Phe Thr Glu Leu Val Cys Ser Leu Ser Asp Ile	670	675	680	2247
15	TAT CAG ATG GAG AAG GAC ATT GCC ATA GAG CAA GAA AGG AAT GCT CGC Tyr Gln Met Glu Lys Asp Ile Ala Ile Glu Gln Glu Arg Asn Ala Arg	685	690	695	2295
20	TAC CGA CCC CCT AAA ATA TTG GAG CCT ACT ACC TTT CAG GAA CCC CCA Tyr Arg Pro Pro Lys Ile Leu Glu Pro Thr Thr Phe Gln Glu Pro Pro	700	705	710	2343
25	CCC AAG CCC AGC CGG CCC AAG TAC AGA CCT CCT CCA CAG ACC AAC CTG Pro Lys Pro Ser Arg Pro Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu	715	720	725	2391
30	CTG GCT CCT AAG CTG CAG TTC CAG GTC CCT GAG GGT CTG TGT GCC AGC Leu Ala Pro Lys Leu Gln Phe Gln Val Pro Glu Gly Leu Cys Ala Ser	735	740	745	2439
35	TCT CCT ACG CTT ACC AGC CCT ATG GAG TAT CCA TCT CCA GTT AAC TCG Ser Pro Thr Leu Thr Ser Pro Met Glu Tyr Pro Ser Pro Val Asn Ser	750	755	760	2487
40	CTG CAC ACC CCA CCT CTC CAC CGG CAC AAT GTC TTC AAG CGC CAC AGC Leu His Thr Pro Pro Leu His Arg His Asn Val Phe Lys Arg His Ser	765	770	775	2535
45	ATG CGG GAG GAG GAC TTC ATC CGG CCC AGT AGC CGA GAA GAG GCC CAG Met Arg Glu Glu Asp Phe Ile Arg Pro Ser Ser Arg Glu Glu Ala Gln	780	785	790	2583
50	CAG CTC TGG GAG GCA GAG AAG ATC AAG ATG AAG CAG GTC CTA GAA AGA Gln Leu Trp Glu Ala Glu Lys Ile Lys Met Lys Gln Val Leu Glu Arg	795	800	805	2631
55	CAG CAG AAG CAG ATG GTG GAA GAT TCC CAG TGG CTG AGG CGA GAG GAA Gln Gln Lys Gln Met Val Glu Asp Ser Gln Trp Leu Arg Arg Glu Glu	815	820	825	2679
60	AGA TGC TTG GAC CCT ATG GTT TAT ATG AAT GAC AAG TCC CCA CTG ACT Arg Cys Leu Asp Pro Met Val Tyr Met Asn Asp Lys Ser Pro Leu Thr	830	835	840	2727
65	CCA GAG AAG GAG GCC GGC TAC ACG GAG TTC ACA GGG CCC CCA CAG AAA Pro Glu Lys Glu Ala Gly Tyr Thr Glu Phe Thr Gly Pro Pro Gln Lys	845	850	855	2775
70	CCA CCT CGG CTC GGT GCA CAG TCC ATT CAG CCC ACA GCC AAC CTG GAC Pro Pro Arg Leu Gly Ala Gln Ser Ile Gln Pro Thr Ala Asn Leu Asp				2823

	860	865	870	
5	AGG ACC GAT GAC CTC GTG TAC CAC AAT GTC ATG ACC CTG GTG GAG GCT Arg Thr Asp Asp Leu Val Tyr His Asn Val Met Thr Leu Val Glu Ala 875 880 885 890	2871		
10	GTG CTG GGA CTC AAG AAC AAG CTT GGC CAG TTG CCC CCT GAG GAC TAT Val Leu Gly Leu Lys Asn Lys Leu Gly Gln Leu Pro Pro Glu Asp Tyr 895 900 905	2919		
15	GTG GTG GTG GTG AAG AAC GTG GGG CTG AAC CTG CGG AAG CTC ATC GGC Val Val Val Val Lys Asn Val Gly Leu Asn Leu Arg Lys Leu Ile Gly 910 915 920	2967		
20	AGT GTG GAC GAT CTC TTG CCC TCC TTG CCG GCA TCT TCG AGG ACA GAG Ser Val Asp Asp Leu Leu Pro Ser Leu Pro Ala Ser Ser Arg Thr Glu 925 930 935	3015		
25	ATT GAA GGG ACC CAG AAA CTG CTC AAC AAA GAC CTG GCA GAG CTC ATC Ile Glu Gly Thr Gln Lys Leu Leu Asn Lys Asp Leu Ala Glu Leu Ile 940 945 950	3063		
30	AAC AAG ATG AAG TTG GCT CAG CAG AAC GCC GTG ACG TCC CTG AGT GAG Asn Lys Met Lys Leu Ala Gln Gln Asn Ala Val Thr Ser Leu Ser Glu 955 960 965 970	3111		
35	GAC TGC AAG CGG CAG ATG CTC ACA GCG TCC CAT ACC CTG GCT GTG GAT Asp Cys Lys Arg Gln Met Leu Thr Ala Ser His Thr Leu Ala Val Asp 975 980 985	3159		
40	GCC AAG AAC CTG CTG GAT GCT GTG GAC CAA GCC AAG GTT GTG GCT AAT Ala Lys Asn Leu Leu Asp Ala Val Asp Gln Ala Lys Val Val Ala Asn 990 995 1000	3207		
45	CTG GCC CAC CCG CCT GCA GAG T GATCAAGAGA GGGGCCACCT GCCTGCATCT Leu Ala His Pro Pro Ala Glu 1005	3259		
50	TCTGCCCCCA CCTGTCTTGG CATACTTTC CTGCCTTGCC TTTGGTTATT GGTCTTCCAG GGAAAGCTGA GAAGAGTCCA TCCCCCTTGC CACTTTGCAC GACGCCCCCT CTTCCCCCAA CCCATCCCAG ACTGTGCTAC TCAGGCTGCA TCTGGACAGA AAGGACTCTG GGCACAGACA CGGGGTGGGG TGACATAGTT CATAGGGGTA CTACTGCCAG CCACTCCCTC TTACCCAGC CTGGGTTGCT GGAGCATCAT TGGGGTCATG AGTGTACCCC TAACGGCCAA GATGGCTTTC TGCATGGACA TTTGAGAGCC AGTATTCCTC CTTCTCTTTC AGCCCTCAGG GACCCCTGAT ACAGAGGGGA CAGAGAGGGG TTTTATTTGT AGAAAAGCTG TGACATGAGG GCTGGACCTG GCTCTCTTGT ACAGTGTACA TTGGAATTTA TTTAATGTGA GTTTGACCTG GATGGACAGC CAAGGGCCAT AGTCCAGGAG CAAACCAATC CAGTCACAGG ACTCTGTGTT TTTATGGAAC	3319 3379 3439 3499 3559 3619 3679 3739 3799		

- 145 -

TGAGTGCCAC AGGGAAGAAG CAGAGAGTCG GAGGTCAGAA TGGGACTTTG TGCCCTTCCT 3859  
 GCGTTTCTCT TCTCCCTCTT TCCTCTCCCC TCTTTTCTTA CGTCTCCTTT TTCTCCTCCC 3919  
 5 CCTTTTCACA TCTGCTCCCC TCCTCTCTCA TGTCTGTGGA GAACATTTAC CTTCTTCTT 3979  
 TTTGATCGGT GGTGAATTA AAATTATTAC CATTTGCTTT GTGAAAAAAA 4029

10 (2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1009 amino acids

(B) TYPE: amino acid

15 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

20 Met Ser Gly Val Ser Glu Pro Leu Ser Arg Val Lys Val Gly Thr Leu  
 1 5 10 15  
 25 Arg Arg Pro Glu Gly Pro Pro Glu Pro Met Val Val Val Pro Val Asp  
 20 25 30  
 Val Glu Lys Glu Asp Val Arg Ile Leu Lys Val Cys Phe Tyr Ser Asn  
 35 40 45  
 30 Ser Phe Asn Pro Gly Lys Asn Phe Lys Leu Val Lys Cys Thr Val Gln  
 50 55 60  
 Thr Glu Ile Gln Glu Ile Ile Thr Ser Ile Leu Leu Ser Gly Arg Ile  
 65 70 75 80  
 35 Gly Pro Asn Ile Gln Leu Ala Glu Cys Tyr Gly Leu Arg Leu Lys His  
 85 90 95  
 40 Met Lys Ser Asp Glu Ile His Trp Leu His Pro Gln Met Thr Val Gly  
 100 105 110  
 Glu Val Gln Asp Lys Tyr Glu Cys Leu His Val Glu Ala Glu Trp Arg  
 115 120 125  
 45 Tyr Asp Leu Gln Ile Arg Tyr Leu Pro Glu Asp Phe Met Glu Ser Leu  
 130 135 140  
 Lys Glu Asp Arg Thr Thr Leu Leu Tyr Phe Tyr Gln Gln Leu Arg Asn  
 145 150 155 160  
 50 Asp Tyr Met Gln Arg Tyr Ala Ser Lys Val Ser Glu Gly Met Ala Leu  
 165 170 175  
 55 Gln Leu Gly Cys Leu Glu Leu Arg Arg Phe Phe Lys Asp Met Pro His  
 180 185 190

- 146 -

Asn Ala Leu Asp Lys Lys Ser Asn Phe Glu Leu Leu Glu Lys Glu Val  
 195 200 205

5 Gly Leu Asp Leu Phe Phe Pro Lys Gln Met Gln Glu Asn Leu Lys Pro  
 210 215 220

Lys Gln Phe Arg Lys Met Ile Gln Gln Thr Phe Gln Gln Tyr Ala Ser  
 225 230 235 240

10 Leu Arg Glu Glu Glu Cys Val Met Lys Phe Phe Asn Thr Leu Ala Gly  
 245 250 255

Phe Ala Asn Ile Asp Gln Glu Thr Tyr Arg Cys Glu Leu Ile Gln Gly  
 260 265 270

15 Trp Asn Ile Thr Val Asp Leu Val Ile Gly Pro Lys Gly Ile Arg Gln  
 275 280 285

Leu Thr Ser Gln Asp Thr Lys Pro Thr Cys Leu Ala Glu Phe Lys Gln  
 290 295 300

Ile Arg Ser Ile Arg Cys Leu Pro Leu Glu Glu Thr Gln Ala Val Leu  
 305 310 315 320

25 Gln Leu Gly Ile Glu Gly Ala Pro Gln Ser Leu Ser Ile Lys Thr Ser  
 325 330 335

Ser Leu Ala Glu Ala Glu Asn Met Ala Asp Leu Ile Asp Gly Tyr Cys  
 340 345 350

30 Arg Leu Gln Gly Glu His Lys Gly Ser Leu Ile Met His Ala Lys Lys  
 355 360 365

Asp Gly Glu Lys Arg Asn Ser Leu Pro Gln Ile Pro Thr Leu Asn Leu  
 370 375 380

Glu Ala Arg Arg Ser His Leu Ser Glu Ser Cys Ser Ile Glu Ser Asp  
 385 390 395 400

40 Ile Tyr Ala Glu Ile Pro Asp Glu Thr Leu Arg Arg Pro Gly Gly Pro  
 405 410 415

Gln Tyr Gly Val Ala Arg Glu Glu Val Val Leu Asn Arg Ile Leu Gly  
 420 425 430

45 Glu Gly Phe Phe Gly Glu Val Tyr Glu Gly Val Tyr Thr Asn His Lys  
 435 440 445

Gly Glu Lys Ile Asn Val Ala Val Lys Thr Cys Lys Lys Asp Cys Thr  
 450 455 460

Gln Asp Asn Lys Glu Lys Phe Met Ser Glu Ala Val Ile Met Lys Asn  
 465 470 475 480

55 Leu Asp His Pro His Ile Val Lys Leu Ile Gly Ile Ile Glu Glu Glu  
 485 490 495



Pro Thr Trp Ile Ile Met Glu Leu Tyr Pro Tyr Gly Glu Leu Gly His  
                   500                                  505                                  510

5 Tyr Leu Glu Arg Asn Lys Asn Ser Leu Lys Val Pro Thr Leu Val Leu  
                   515                                  520                                  525

Tyr Thr Leu Gln Ile Cys Lys Ala Met Ala Tyr Leu Glu Ser Ile Asn  
                   530                                  535                                  540

10 Cys Val His Arg Asp Ile Ala Val Arg Asn Ile Leu Val Ala Ser Pro  
                   545                                  550                                  555                                  560

15 Glu Cys Val Lys Leu Gly Asp Phe Gly Leu Ser Arg Tyr Ile Glu Asp  
                                   565                                  570                                  575

Glu Asp Tyr Tyr Lys Ala Ser Val Thr Arg Leu Pro Ile Lys Trp Met  
                   580                                  585                                  590

20 Ser Pro Glu Ser Ile Asn Phe Arg Arg Phe Thr Thr Ala Ser Asp Val  
                   595                                  600                                  605

Trp Met Phe Ala Val Cys Met Trp Glu Ile Leu Ser Phe Gly Lys Gln  
                   610                                  615                                  620

25 Pro Phe Phe Trp Leu Glu Asn Lys Asp Val Ile Gly Val Leu Glu Lys  
                   625                                  630                                  635                                  640

30 Gly Asp Arg Leu Pro Lys Pro Glu Leu Cys Pro Pro Val Leu Tyr Thr  
                                   645                                  650                                  655

Leu Met Thr Arg Cys Trp Asp Tyr Asp Pro Ser Asp Arg Pro Arg Phe  
                   660                                  665                                  670

35 Thr Glu Leu Val Cys Ser Leu Ser Asp Ile Tyr Gln Met Glu Lys Asp  
                   675                                  680                                  685

Ile Ala Ile Glu Gln Glu Arg Asn Ala Arg Tyr Arg Pro Pro Lys Ile  
                   690                                  695                                  700

40 Leu Glu Pro Thr Thr Phe Gln Glu Pro Pro Pro Lys Pro Ser Arg Pro  
                   705                                  710                                  715                                  720

Lys Tyr Arg Pro Pro Pro Gln Thr Asn Leu Leu Ala Pro Lys Leu Gln  
                   725                                  730                                  735

Phe Gln Val Pro Glu Gly Leu Cys Ala Ser Ser Pro Thr Leu Thr Ser  
                   740                                  745                                  750

50 Pro Met Glu Tyr Pro Ser Pro Val Asn Ser Leu His Thr Pro Pro Leu  
                   755                                  760                                  765

His Arg His Asn Val Phe Lys Arg His Ser Met Arg Glu Glu Asp Phe  
                   770                                  775                                  780

55 Ile Arg Pro Ser Ser Arg Glu Glu Ala Gln Gln Leu Trp Glu Ala Glu

- 148 -

	785		790		795		800									
	Lys	Ile	Lys	Met	Lys	Gln	Val	Leu	Glu	Arg	Gln	Gln	Lys	Gln	Met	Val
					805					810					815	
5	Glu	Asp	Ser	Gln	Trp	Leu	Arg	Arg	Glu	Glu	Arg	Cys	Leu	Asp	Pro	Met
					820					825					830	
10	Val	Tyr	Met	Asn	Asp	Lys	Ser	Pro	Leu	Thr	Pro	Glu	Lys	Glu	Ala	Gly
					835					840					845	
	Tyr	Thr	Glu	Phe	Thr	Gly	Pro	Pro	Gln	Lys	Pro	Pro	Arg	Leu	Gly	Ala
					850					855					860	
15	Gln	Ser	Ile	Gln	Pro	Thr	Ala	Asn	Leu	Asp	Arg	Thr	Asp	Asp	Leu	Val
					865					870					875	
	Tyr	His	Asn	Val	Met	Thr	Leu	Val	Glu	Ala	Val	Leu	Gly	Leu	Lys	Asn
					885					890					895	
20	Lys	Leu	Gly	Gln	Leu	Pro	Pro	Glu	Asp	Tyr	Val	Val	Val	Val	Lys	Asn
					900					905					910	
25	Val	Gly	Leu	Asn	Leu	Arg	Lys	Leu	Ile	Gly	Ser	Val	Asp	Asp	Leu	Leu
					915					920					925	
	Pro	Ser	Leu	Pro	Ala	Ser	Ser	Arg	Thr	Glu	Ile	Glu	Gly	Thr	Gln	Lys
					930					935					940	
30	Leu	Leu	Asn	Lys	Asp	Leu	Ala	Glu	Leu	Ile	Asn	Lys	Met	Lys	Leu	Ala
					945					950					955	
	Gln	Gln	Asn	Ala	Val	Thr	Ser	Leu	Ser	Glu	Asp	Cys	Lys	Arg	Gln	Met
					965					970					975	
35	Leu	Thr	Ala	Ser	His	Thr	Leu	Ala	Val	Asp	Ala	Lys	Asn	Leu	Leu	Asp
					980					985					990	
40	Ala	Val	Asp	Gln	Ala	Lys	Val	Val	Ala	Asn	Leu	Ala	His	Pro	Pro	Ala
					995					1000					1005	

Glu

45 (2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

55

## (ix) FEATURE:

(A) NAME/KEY: Peptide

(B) LOCATION: 6

5 (D) OTHER INFORMATION: /label= OTHER FEATURE  
/note= "X=F OR Y"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

10 Ser Asp Val Trp Ser Xaa Gly  
1 5

## (2) INFORMATION FOR SEQ ID NO:6:

15 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
20 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

25 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

SWRTCNACCC ANSWRWANCC

20

30 (2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 6 amino acids  
(B) TYPE: amino acid  
35 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

40

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

45 Asp Leu Ala Ala Arg Asn  
1 5

## (2) INFORMATION FOR SEQ ID NO:8:

50 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 18 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

55 (ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

5 CGACGAYCTN GCNRCNAA

18

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

10 (A) LENGTH: 5 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

15

(v) FRAGMENT TYPE: internal

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Trp Met Ala Pro Glu  
1 5

25 (2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

30 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

40 GTACCYTCVG GNGCCATCCA

20

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

45 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

55 CGGGCCGTGC TGGAGCTCAA

20

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

15 GTCCGTGAAG ATGACGGCAA

20

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 21 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## 25 (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

30 AAAGCTGTCA TCGAGATGTC C

21

## (2) INFORMATION FOR SEQ ID NO:14:

## 35 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

40

## (ii) MOLECULE TYPE: cDNA

## 45 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TCGGTGGGTG CTGGCTGGTA GG

22

## (2) INFORMATION FOR SEQ ID NO:15:

50

## (i) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 32 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

- 152 -

(ii) MOLECULE TYPE: cDNA

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

ATCTGGCACC ACACCTTCTA CAATGAGCTG CG

32

(2) INFORMATION FOR SEQ ID NO:16:

10

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

15

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGTCATACTC CTGCTTGCTG ATCCACATCT GC

32

25 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

30

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

AGCTTATGGA CTACAAGGAC GACGATGACA GGGG

34

40

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

45

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

55 AATTCCCTTG TCATCGTCGT CCTTATGGTC CATA

34

What is claimed is:

- 5 1. An isolated *RAFTK* nucleic acid molecule from a vertebrate organism.
2. The isolated nucleic acid molecule of claim 1 having a nucleic acid sequence shown in one of SEQ ID NOs:1 or 3, or a complement or fragment thereof.
- 10 3. The isolated nucleic acid molecule of claim 2 which has at least 70% homologous to a nucleic acid molecule shown in one of SEQ ID NOs:1 or 3.
4. The isolated nucleic acid molecule of claim 1, which encodes a polypeptide with a *RAFTK* bioactivity.
- 15 5. An isolated nucleic acid molecule of claim 1, which encodes a polypeptide shown in SEQ ID NO. 2 or 4.
6. An isolated nucleic acid molecule of claim 1, which is capable of hybridizing to a nucleic acid molecule of one of SEQ ID NOs:1 or 3 under stringent conditions.
- 20 7. An isolated nucleic acid molecule of claim 1, which encodes a polypeptide that is at least 70% identical to the polypeptide shown in SEQ ID NOs.: 2 or 4.
- 25 8. An isolated nucleic acid molecule of claim 2, which comprises the coding region of one of SEQ ID NOs:1 or 3.
9. An isolated nucleic acid molecule of claim 1, which hybridizes to at least 6 consecutive nucleotides of the *RAFTK* gene shown in one of one of SEQ ID NOs:1 or 3.
- 30 10. An isolated nucleic acid molecule of claim 9, which further comprises a label.
- 35 11. An expression vector, comprising a nucleic acid molecule of claim 2 operably linked to a transcriptional regulatory sequence.

12. An expression vector of claim 11, which is capable of replicating in a cell.
- 5 13. A host cell transfected with an expression vector of claim 12.
14. A method of making a *RAFTK* polypeptide comprising the steps of:
- a. culturing the cell of claim 13 in an appropriate culture medium to produce a *RAFTK* polypeptide; and
- 10 style="padding-left: 40px;">b. isolating the *RAFTK* polypeptide.
15. A transgenic animal in which expression of a genomic sequence encoding a functional *RAFTK* polypeptide is enhanced, induced, prevented or suppressed.
- 15 16. An isolated polypeptide of a vertebrate organism having a *RAFTK* bioactivity.
17. A polypeptide of claim 16, which is at least 70% homologous to the polypeptide shown in SEQ ID NOs. 2 or 4.
- 20 18. A polypeptide of claim 17, which has a molecular weight of approximately 123kD.
- 25 19. A fusion protein comprising a polypeptide of claim 17 and a second polypeptide, said fusion protein containing a detectable label or a matrix binding domain.
- 30 20. A pharmaceutical preparation comprising a therapeutically effective amount of the polypeptide of claim 17 and a pharmaceutically acceptable carrier.
21. An antibody which is specifically reactive with an epitope of the polypeptide of claim 16.
- 35 22. A method for modulating one or more of growth, differentiation, hematopoiesis, or survival in a cell, comprising treating the cell with an effective amount of an agent which modulates the activity of a *RAFTK* protein thereby altering,



relative to the cell in the absence of the agent, at least one of (i) rate of growth, (ii) differentiation, (iii) hematopoiesis or (iv) survival of the cell.

23. The method of claim 22, wherein the cell is selected from the group  
5 consisting of a mast cell, a melanocyte, and a megakaryocyte.

24. A method for modulating one or more of cell adhesion, migration,  
phagocytosis, or motility of a cell, comprising treating the cell with an effective  
amount of an agent which modulates the activity of a *RAFTK* protein thereby altering,  
10 relative to the cell in the absence of the agent, at least one of (i) cell adhesion, (ii)  
migration, (iii) phagocytosis, or (iv) motility of the cell.

25. A method of claim 24, wherein said cell adhesion is modulated by  
modulating focal adhesion formation in a cell, said method comprising treating the  
15 cell with an effective amount of an agent which modulates the activity of a *RAFTK*  
protein.

26. The method of claim 25, wherein said method is used to treat  
metastasis by a tumor cell.  
20

27. A diagnostic assay for identifying a cell or cells at risk for a disorder  
characterized by unwanted cell proliferation or differentiation, comprising detecting,  
in a cell sample, the presence or absence of a genetic lesion characterized by at least  
one of (i) aberrant modification or mutation of a gene encoding a *RAFTK* protein, and  
25 (ii) mis-expression of said gene; wherein a wild-type form of said gene encodes an  
polypeptide with a *RAFTK* bioactivity.

28. The assay of claim 27, wherein detecting said lesion includes:  
a. providing a diagnostic probe comprising a nucleic acid  
30 including a region of nucleotide sequence which hybridizes to a sense or antisense  
sequence of said gene, or naturally occurring mutants thereof, or 5' or 3' flanking  
sequences naturally associated with said gene;  
b. combining said probe with nucleic acid of said cell sample; and  
c. detecting, by hybridization of said probe to said cellular nucleic  
35 acid, the existence of at least one of a deletion of one or more nucleotides from said  
gene, an addition of one or more nucleotides to said gene, a substitution of one or  
more nucleotides of said gene, a gross chromosomal rearrangement of all or a portion

of said gene, a gross alteration in the level of an mRNA transcript of said gene, or a non-wild type splicing pattern of an mRNA transcript of said gene.

29. A method of preparing differentiated blood cells comprising  
5 modulating the activity of a *RAFTK* protein in a progenitor stem cell.

30. The method of claim 29, wherein the differentiated blood cells are megakaryocytes.

10 31. A population of megakaryocytes prepared using the method of claim 30.

32. A population of platelets prepared using the method of claim 30.

15 33. An assay for screening test compounds to identify compounds which modulate *RAFTK* interaction with cellular proteins, comprising:

- a. providing a reaction mixture including a *RAFTK* protein, a *RAFTK*-binding protein, and a test compound; and
- b. detecting the interaction of the *RAFTK* protein and the *RAFTK*-  
20 binding protein,

wherein a statistically significant change in the interaction of the proteins in the presence of the test compound is indicative of the capability of a compound to modulate a bioactivity of a *RAFTK* polypeptide.

25 34. The assay of claim 33, wherein the *RAFTK*-binding protein is selected from the group consisting of paxillin, protein kinase C- $\alpha$ , Protein kinase C- $\delta$ , src, fyn, Grb2, PI3 kinase, and the c-fms receptor, and calpain.

30 35. The assay of claim 33, wherein the detecting step comprises determining the level of phosphorylation of *RAFTK* or the *RAFTK* binding protein.

36. The assay of claim 33, wherein the reaction mixture is selected from the group consisting of a reconstituted protein mixture and a cell lysate.

37. The assay of claim 33, wherein the *RAFTK* protein is a recombinant protein.

38. The assay of claim 33, wherein one or both of the *RAFTK* protein and  
5 *RAFTK*-binding protein is a fusion protein.

39. The assay of claim 33, wherein at least one of the *RAFTK* protein and  
*RAFTK*-binding protein comprises an endogenous detectable label for detecting the  
formation of said complex.

10

40. The method of claim 33, which reaction mixture is a whole cell, and  
interaction of the *RAFTK* protein and *RAFTK*-binding protein is detected in a two  
hybrid assay system.

15

41. A *RAFTK* inhibitor identified using the assay of claim 33.

42. A pharmaceutical preparation comprising (i) the composition of claim  
41 in an amount effective for inhibiting proliferation of a cell, and (ii) a  
pharmaceutically acceptable diluent.

20

43. A method for modulating one or more of growth, differentiation, or  
survival of a megakaryocytic cell, comprising treating the cell with an effective  
amount of the preparation of claim 42 so as to modulate *RAFTK* activity and alter,  
relative to the cell in the absence of the agent, at least one of (i) the growth, (ii)  
25 migration, (iii) differentiation state, or (iv) survival of the cell.

1/5

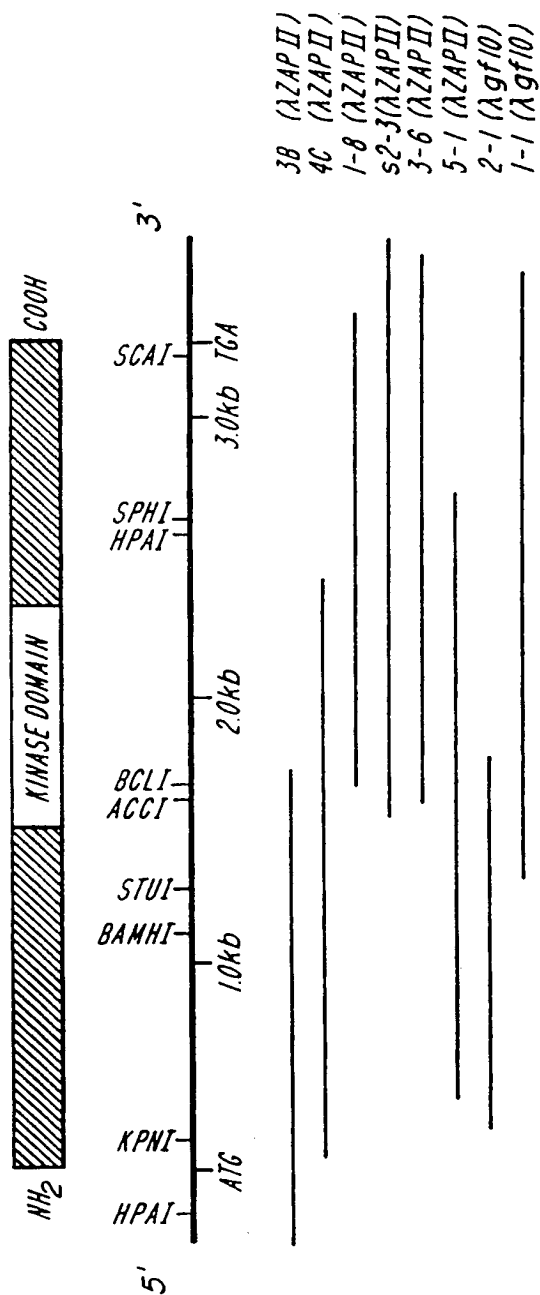


FIG. 1

2/5

[illegible]

FIG. 2

[illegible]

FIG. 3



## BXD RI LINE DATA

	<i>Gnrh</i>	<i>Raftk</i>
01	B	B
02	D	D
05	B	B
06	D	D
08	D	D
09	D	D
11	D	D
12	B	B
13	B	B
14	D	D
15	D	D
16	D	D
18	B	B
19	D	D
20	B	B
21	B	B
22	D	D
23	B	B
24	B	B
25	D	D
27	B	B
28	B	B
29	D	D
30	B	B
31	B	B
32	D	D

*Gnrh* --- < *lcM* --- *Raftk*  
95 % confidence limits  
0-4.1 Map Units

***FIG. 5***



# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/14093

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C12N 15/52, 15/63, 9/00; C07K 19/00; C12Q 1/68; A61K 38/43;  
US CL :536/23.5, 24.31; 435/6, 69.1, 320.1, 325; 530/350; 514/12  
According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.5, 24.31; 435/6, 69.1, 320.1, 325; 530/350; 514/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

STN/Medline, HCaPlus; APS/USPAT

search terms: RAFTK, related adhesion focal tyrosine kinase, PYK2, CAK beta

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X ----- Y	LEV, S. et al. Protein Tyrosine Kinase PYK2 Involved in Ca <sup>2+</sup> -Induced Regulation of Ion Channel and MAP Kinase Functions. Nature. 31 August 1995. Volume 376, pages 737-745, especially pages 738-740.	1-14, 16-20 ----- 27-28
X ----- Y	AVRAHAM, S. et al. Identification and Characterization of a Novel Related Adhesion Focal Tyrosine Kinase (RAFTK) from Megakaryocytes and Brain. Journal of Biological Chemistry. 17 November 1995. Volume 270, Number 46, pages 27742-27751, especially pages 27742-27747.	1-14, 16-20 ----- 27-28

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	*X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	*Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	* & * document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means	
*P* document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

14 OCTOBER 1997

Date of mailing of the international search report

06 JAN 1998

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized Officer

KAREN E. BROWN

Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/US97/14093

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SASAKI, H. et al. Cloning and Characterization of Cell Adhesion Kinase Beta, a Novel Protein-Tyrosine Kinase of the Focal Adhesion Kinase Subfamily. Journal of Biological Chemistry. 08 September 1995. Volume 270, Number 36, pages 21206-21219, especially pages 21206-21211.	1-14, 16-20
X	LI, J. et al. Characterization of RAFTK, a Novel Focal Adhesion Kinase, and Its Integrin-Dependent Phosphorylation and Activation in Megakaryocytes. Blood. 15 July 1996. Volume 88, Number 2, pages 417-428, especially pages 418-427.	1-14, 16-20

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/14093

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
  
2. ☒ Claims Nos.: 1-14, 17-20 and 27-28  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:  
  
Please See Extra Sheet.
  
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
  
4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:  
1-14, 16-20, 27-28

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/14093

## BOX I. OBSERVATIONS WHERE CLAIMS WERE FOUND UNSEARCHABLE

### 2. Where no meaningful search could be carried out, specifically:

All of the claims are unsearchable to the extent that they require reference to the specified sequences from the sequence listing. Because the sequence listing supplied by the Applicant was unreadable, no meaningful search of the sequences per se can be carried out by this Authority. However, the subject matter of the claims has been searched to the extent possible with reference to the balance of the description.

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claim(s) 1-14, 16-20 and 27-28, drawn to RAFTK polynucleotides, expression vectors, host cells, a method of producing RAFTK recombinantly, RAFTK polypeptide and a method of using RAFTK polynucleotides in a diagnostic assay.

Group II, claim(s) 15, drawn to a transgenic animal in which RAFTK activity is altered.

Group III, claim(s) 21, drawn to an anti-RAFTK antibody.

Group IV, claim(s) 22-26, 29-32 and 43, drawn to a method of altering RAFTK activity using an agent which modulates the effect of RAFTK protein.

Group V, claim(s) 33-40, drawn to a method of using RAFTK protein to identify compounds which modulate RAFTK activity.

Group VI, claim(s) 41-42, drawn to RAFTK inhibitors.

The inventions listed as Groups I-VI do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

The special technical feature of Group I is a polynucleotide encoding RAFTK.

Group II does not share the same special technical feature as Group I because the the polynucleotide of Group I is not required for the production of the transgenic animal of Group II.

Group III does not share the same special technical feature as Group I because the polynucleotide of Group I does not encode the antibody of Group III.

Group IV does not share the same special technical feature of Group I because the method of Group IV neither uses nor produces the polynucleotide of Group I, and the polynucleotide of Group I is not required in any steps of the method of Group IV.

Group V does not share the same special technical feature of Group I because the method of Group V neither uses nor produces the polynucleotide of Group I, and the polynucleotide of Group I is not required in any steps of the method of Group V.

Group VI does not share the same special technical feature of Group I because the polynucleotide of Group I is not an inhibitor of Group VI nor does it encode the inhibitor of Group VI, and thus the inhibitor of Group VI does not share the same special technical feature as the polynucleotide of Group I.

Accordingly, the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive concept.